Cours d'algèbre M1

Pierre Guillot

4 septembre 2023

Table des matières

Chapitre 1

Groupes & Anneaux

§1.1 Groupes

Nous étudierons les groupes en détail dans un chapitre ultérieur, mais ici nous rappelons simplement les définitions.

Définitions & Premiers exemples

Définition 1. Un *groupe* est un ensemble G, sur lequel on a une opération

$$\begin{aligned} G \times G &\longrightarrow G \\ (g, h) &\mapsto g \star h \,. \end{aligned}$$

On suppose aussi qu'il y a un élément distingué $e \in G$, appelé *l'élément neutre*. Enfin, à tout $g \in G$ on associe un certain élément $\tilde{g} \in G$. On exige que les conditions suivantes soient satisfaites, pour tous les $g, h, k \in G$:

1. $(g \star h) \star k = g \star (h \star k)$ (associativité)
2. $e \star g = g \star e = g$
3. $g \star \tilde{g} = \tilde{g} \star g = e$.

Exemple 2. Prenons

$$G = \{A \in \mathrm{M}_n(\mathbb{K}) \mid A^{-1} \text{ existe}\} \,.$$

(Ici, et partout ailleurs, $\mathbb{K}$ est un corps ; mais comme nous n'avons pas encore fait de rappels sur les corps, disons que $\mathbb{K} = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$ dans un premier

temps, si on veut.) Notre loi sera $A \star B = AB$, le produit usuel des matrices ; l'élément neutre que nous choisissons est $e = I$, la matrice identité ; et enfin on prend $\tilde{A} = A^{-1}$. Les conditions sont évidemment remplies. Donc G, avec ces choix, est un groupe ; on le note en général $GL_n(\mathbb{K})$, où GL est là pour « groupe linéaire ».

Exemple 3. Soit X un ensemble, et posons

$$G = \{f : X \to X \mid f \text{ est une bijection}\}.$$

On prend $f \star g = f \circ g$, la composition ; on prend $e = Id$, la fonction identité, c'est-à-dire $Id(x) = x$; et enfin, $\tilde{f} = f^{-1}$, la réciproque de f, qui est aussi une bijection (et donc est bien dans G). Alors G est un groupe. On le note souvent $S(X)$, le *groupe symétrique de* X. Lorsque $X = \{1, 2, \ldots, n\}$, on note simplement S_n (ou Σ_n, ou $\mathfrak{S}_n$).

Exemple 4. Prenons $G = \mathbb{Z}$, $n \star m = n+m$ (l'addition usuelle), $e = 0$, et $\tilde{n} = -n$. Alors G est aussi un groupe !

Définition 5. Soit G un groupe. On dit que G est *abélien*, ou *commutatif*, lorsque $g \star h = h \star g$ pour tous les $g, h \in G$.

Exemple 6. Le groupe $\mathbb{Z}$ (pour l'addition) est abélien. Mais pour $n \geq 2$, le groupe $GL_n(\mathbb{K})$ n'est pas abélien.

On va simplifier un peu les notations. Nous dirons que le groupe G est « en notation multiplicative » lorsque son opération est notée gh au lieu de $g \star h$, son élément neutre est noté 1, et on écrit g^{-1} au lieu de $\tilde{g}$. Par défaut, les groupes seront tous en notation multiplicative.

Lorsqu'un groupe est abélien, et seulement dans ce cas, on s'autorise parfois à le mettre « en notation additive », c'est-à-dire qu'on écrit $g+h$ pour $g \star h$, on écrit 0 pour l'élément neutre e, et on écrit $-x$ pour $\tilde{x}$ (ainsi que $x - y$ pour $x + (-y)$).

Remarque. Voici des rappels très, très brefs sur ce qu'est $\mathbb{Z}/n\mathbb{Z}$, à toutes fins utiles (vous êtes censés connaître déjà, mais ajoutons ici que dans un prochain chapitre, nous étudierons les « quotients » en général, ce qui donnera une

nouvelle définition de $\mathbb{Z}/n\mathbb{Z}$). Fixons donc un entier $n \geq 1$. Pour tout $x \in \mathbb{Z}$, on va noter $\overline{x}$ le reste dans la division euclidienne de x par n (et $\overline{x}$ est parfois prononcé « x modulo n »). Pour tout x on a

$$\overline{x} \in \{0, 1, 2, \ldots, n-1\} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\} =: \mathbb{Z}/n\mathbb{Z}\,.$$

On montre que $\mathbb{Z}/n\mathbb{Z}$ est un groupe abélien, naturellement en notation additive ici, et que $\overline{x+y} = \overline{x} + \overline{y}$ (l'élément neutre est bien sûr $\overline{0}$, et $-\overline{x} = \overline{-x}$).

Sous-groupes

Les exemples de groupes seront en général obtenus par les considérations suivantes.

Définition 7. Soit G un groupe (en notation multiplicative), et soit $H \subset G$. On dit que H est un *sous-groupe* de G lorsque

1. $1 \in H$;
2. si $g \in H$ et $h \in H$, alors $gh \in H$;
3. si $g \in H$, alors $g^{-1} \in H$.

Remarque. Une observation évidente, mais fondamentale : si H est un sous-groupe de G, alors H est lui-même un groupe !

Exemple 8. Soit $n \geq 1$, et posons

$$\mu_n = \mu_n(\mathbb{C}) = \{z \in \mathbb{C}^* \mid z^n = 1\}\,.$$

Alors μ_n est un sous-groupe de $\mathbb{C}^*$. En effet, si $z^n = w^n = 1$, alors $(zw)^n = z^n w^n = 1$, et $(z^{-1})^n = (1/z)^n = 1/z^n = 1$. Et bien sûr $1 \in \mu_n$.

Exemple 9. Soit

$$T = \left\{ \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \mid a, b, c \in \mathbb{R},\ ab \neq 0 \right\}\,.$$

Vérifions que T est un sous-groupe de $GL_2(\mathbb{R})$; on note déjà que $T \subset GL_2(\mathbb{R})$, car ab est le déterminant de la matrice proposée. On calcule donc

$$\begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \begin{pmatrix} a' & c' \\ 0 & b' \end{pmatrix} = \begin{pmatrix} aa' & ac' + cb' \\ 0 & bb' \end{pmatrix} \in T$$

et

$$\begin{pmatrix} a & c \\ 0 & b \end{pmatrix}^{-1} = \frac{1}{ab}\begin{pmatrix} b & -cc \\ 0 & a \end{pmatrix} = \begin{pmatrix} \frac{1}{a} & -\frac{c}{ab} \\ 0 & \frac{1}{b} \end{pmatrix} \in T\,,$$

et la dernière vérification est évidente.

Définition 10. Soit G un groupe en notation multiplicative, soit $g \in G$, et soit $n \in \mathbb{N}^*$. On note g^n pour $gg\cdots g$, répété n fois ; pour $n \in \mathbb{Z}$, mais $n < 0$, on définit $g^n = g^{-1}g^{-1}\cdots g^{-1}$, répété $-n$ fois ; et on pose $g^0 = 1$. Les règles de calcul usuelles s'appliquent sans aucune surprise (on le vérifie), comme $g^n g^m = g^{n+m}$, et $(g^n)^{-1} = g^{-n}$.

Si G est en notation additive, on définit de même ng pour $n \in \mathbb{Z}$. En particulier $0g = 0$.

Définition 11. Soit G un groupe et $g \in G$. Le *sous-groupe cyclique engendré par g*, par définition, est

$$\langle g\rangle = \{g^n \mid n \in \mathbb{Z}\}\,.$$

Si G est en notation additive, on a

$$\langle g\rangle = \{ng \mid n \in \mathbb{Z}\}\,.$$

Pour justifier ce nom, il faut vérifier que $\langle g\rangle$ est bien un sous-groupe de G ! Mais c'est assez évident, ça provient des identités $g^n g^m = g^{n+m}$ etc. Ajoutons une autre définition :

Définition 12. On dit que le groupe G est *cyclique* s'il existe $g \in G$ tel que $G = \langle g\rangle$.

Exemple 13. Le groupe $\mathbb{Z}$ est cyclique, puisque $\mathbb{Z} = \langle 1\rangle$. De même $\mathbb{Z}/n\mathbb{Z} = \langle \overline{1}\rangle$ est cyclique. (Attention, dans les deux cas, la notation est additive). Nous verrons plus loin que tous les groupes cycliques sont « modelés » sur ceux-ci. Pour un exemple en notation multiplicative, considérer μ_n, le groupes des racines n-ièmes dans $\mathbb{C}$: il est cyclique, puisque $\mu_n = \langle \exp(\frac{2i\pi}{n})\rangle$.

Exemple 14. Soit G un groupe ne comportant que deux éléments. Montrons que G est cyclique. En effet, appelons g l'unique élément de G tel que $g \neq 1$. Alors $G = \{g^0 = 1, g\} = \langle g\rangle$.

Exemple 15. Un groupe cyclique est toujours abélien. Donc $GL_2(\mathbb{K})$ n'est pas cyclique, par exemple.

Définition 16. Si G est un groupe fini, alors le nombre d'éléments de G est appelé *l'ordre* de G, noté $|G|$. Si G n'est pas fini, on dit parfois que son ordre est ∞. Si $g \in G$, où G est un groupe quelconque, alors l'ordre du groupe $\langle g \rangle$ est appelé *l'ordre de* g (ça peut donc être ∞).

Lemme 17. *Soit $g \in G$. L'ordre de g est le plus petit entier $k > 0$ tel que $g^k = 1$, ou alors ∞ s'il n'en existe pas. On a alors*

$$\langle g \rangle = \{1, g, g^2, \ldots, g^{k-1}\}.$$

Enfin, si $n, m \in \mathbb{Z}$, alors

$$g^n = g^m \iff n \equiv m \bmod k.$$

En particulier, on voit que $g^n = 1 \Leftrightarrow k$ divise n.

Démonstration. Soit $H = \langle g \rangle$. Supposons d'abord que H est fini, c'est-à-dire que l'ordre de g est fini. Alors les éléments g^i, pour $i \in \mathbb{N}$, ne sont pas tous différents. Soient donc $i < j$ deux entiers tels que $g^i = g^j$. On en déduit $1 = g^{j-i}$, donc il existe des entiers $k > 0$ tels que $g^k = 1$ (par exemple $k = j - i$). Par contraposée, on obtient que, si des entiers tels que k n'existent pas, alors l'ordre de g est ∞.

On suppose donc qu'il existe de tels entiers, et on appelle k le plus petit d'entre eux. Les éléments $1, g, g^2, \ldots, g^{k-1}$ sont tous différents, car $g^i = g^j$ avec $i < j < k$ entraînerait $g^{j-i} = 1$, mais $j - i < k$, c'est absurde. Par ailleurs, prenons $n \in \mathbb{Z}$ et écrivons la division euclidienne $n = kq + r$ avec $0 \leq r < k$. On a alors $g^n = g^{kq+r} = (g^k)^q g^r = 1^q g^r = g^r$. Donc $g^n \in \{1, g, \ldots, g^{k-1}\}$.

Finalement $H = \{1, g, \ldots, g^{k-1}\}$. Il comporte donc k éléments, ce qui montre que l'ordre de g est fini, et même que cet ordre est k. Par ailleurs, on a constaté que g^n ne dépend que de $\overline{n} = r \in \mathbb{Z}/k\mathbb{Z}$, donc le reste est clair. □

Homomorphismes

Définition 18. Soit G_1 un groupe, avec élément neutre e_1 et opération notée $\star$, et soit G_2 un autre groupe, avec élément neutre e_2 et opération notée $\bullet$. Un

homomorphisme de groupes entre G_1 et G_2 est une fonction $\varphi : G_1 \to G_2$ telle que

$$\varphi(g \star h) = \varphi(g) \bullet \varphi(h) ,$$

pour tous les g, h. Si les deux groupes sont en notation multiplicative, cette condition s'écrit

$$\varphi(gh) = \varphi(g)\varphi(h) .$$

Si, par exemple, G_1 est en notation additive, et G_2 en multiplicative, la condition est

$$\varphi(g + h) = \varphi(g)\varphi(h) .$$

On abrège souvent « homomorphisme de groupes » en « homomorphisme » lorsque la confusion n'est pas possible.

Remarque. Lorsque φ est un homomorphisme, on a automatiquement $\varphi(e_1) = e_2$ (ce qui s'écrit $\varphi(1) = 1$ en notation multiplicative). En effet, $e_1^2 = e_1$, donc $\varphi(e_1^2) = \varphi(e_1)^2 = \varphi(e_1)$, ce qui s'écrit $x^2 = x$ en posant $x = \varphi(e_1)$. En multipliant par x^{-1} on obtient $x^{-1}x^2 = x = x^{-1}x = e_2$. Mais de toute façon, dans tous les exemples, la propriété $\varphi(e_1) = e_2$ est toujours évidente. On vous laisse montrer, de la même façon, que l'on a toujours $\varphi(x^{-1}) = \varphi(x)^{-1}$ automatiquement.

Exemple 19. On considère $\mathbb{R}$, qui est un groupe pour l'addition, et $\mathbb{R}^\times$ qui est un groupe pour la multiplication. Alors l'exponentielle est un homomorphisme $\mathbb{R} \to \mathbb{R}^\times$. Son image est le sous-groupe $\mathbb{R}^{>0}$. Le logarithme est un homomorphisme $\mathbb{R}^{>0} \to \mathbb{R}$.

Exemple 20. Reprenons le groupe

$$T = \left\{ \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \mid a, b, c \in \mathbb{R},\ ab \neq 0 \right\} \subset GL_2(\mathbb{R}) .$$

Alors on a un homomorphisme $T \to \mathbb{R}^*$ défini par

$$\begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \mapsto a \qquad (\text{ou } b) .$$

Exemple 21. L'application

$$\det\colon GL_2(\mathbb{K}) \longrightarrow \mathbb{K}^*$$

est un homomorphisme : $\det(AB) = \det(A)\det(B)$.

Exemple 22. Soit G un groupe quelconque, et $g \in G$. Alors on a toujours un homomorphisme

$$\begin{aligned} \varphi_g\colon \mathbb{Z} &\longrightarrow G \\ n &\mapsto g^n\,. \end{aligned}$$

(En effet $\varphi_g(n+m) = g^{n+m} = g^n g^m = \varphi(n)\varphi(m)$.) Nous l'appellerons simplement « l'homomorphisme défini par le choix de $g \in G$ ».

Définition 23. Un *isomorphisme* entre G_1 et G_2 est un homomorphisme $G_1 \to G_2$ qui est aussi une bijection. Lorsqu'il existe au moins un tel isomorphisme, on dit que G_1 et G_2 sont *isomorphes*, et on note $G_1 \cong G_2$.

Lemme 24. *Si $\varphi\colon G_1 \to G_2$ est un isomorphisme, alors la bijection réciproque $\varphi^{-1}\colon G_2 \to G_1$ est aussi un isomorphisme.*

Démonstration. Soient $x, y \in G_2$. Soient $g, h \in G_1$ tels que $\varphi(g) = x$ et $\varphi(h) = y$; en d'autres termes, soient $g = \varphi^{-1}(x)$, et $h = \varphi^{-1}(y)$. Alors $\varphi^{-1}(xy) = \varphi^{-1}(\varphi(g)\varphi(h)) = \varphi^{-1}(\varphi(gh)) = gh$, en utilisant que φ est un homomorphisme. Donc $\varphi^{-1}(xy) = \varphi^{-1}(x)\varphi^{-1}(y)$, ce qui montre bien que φ^{-1} est un homomorphisme (et donc un isomorphisme puisque c'est une bijection). □

Pour illustrer la notion d'isomorphisme, nous allons montrer la chose suivante :

Proposition 25. *Soit $G = \langle g \rangle$ un groupe cyclique. Alors G est soit isomorphe à $\mathbb{Z}$, soit isomorphe à $\mathbb{Z}/k\mathbb{Z}$ pour un certain $k \geq 1$.*

Démonstration. On considère l'homomorphisme $\mathbb{Z} \to G$ défini par g. Envisageons d'abord le cas où φ_g est injective. Elle est surjective par définition (G étant cyclique), donc c'est un isomorphisme, et $G \cong \mathbb{Z}$.

Supposons donc que φ_g n'est pas injective. On retrouve un argument connu : il existe $i < j$ avec $\varphi_g(i) = g^i = \varphi_g(j) = g^j$, donc $g^{j-i} = 1$, avec $j - i > 0$. Par un lemme précédent, l'ordre de g est fini ; appelons-le k. Définissons alors $\psi : \mathbb{Z}/k\mathbb{Z} \to G$ par $\psi(\overline{n}) = g^n$. Le même lemme que précédemment montre que ψ est bien défini, et aussi que c'est une bijection.

Il faut tout de même vérifier que c'est un homomorphisme. Commençons par écrire $\psi(\overline{n})\psi(\overline{m}) = g^n g^m = g^{n+m} = \psi(\overline{n+m})$. Or l'addition de $\mathbb{Z}/k\mathbb{Z}$ vérifie $\overline{n+m} = \overline{n}+\overline{m}$, d'où $\psi(\overline{n})\psi(\overline{m}) = \psi(\overline{n}+\overline{m})$, ce qui achève l'argument. □

Terminons avec la notion de noyau :

Définition 26. Soient G_1 et G_2 deux groupes, et soit $\varphi : G_1 \longrightarrow G_2$ un homomorphisme. Le *noyau* de φ est

$$\ker(\varphi) = \{g \in G_1 \mid \varphi(g) = e_2\}\,.$$

(Bien sûr e_2 est l'élément neutre de G_2.) C'est un sous-groupe de G_1 (petit exo).

Exemple 27. Le noyau du déterminant, qui est un homomorphisme $GL_n(\mathbb{K}) \longrightarrow \mathbb{K}^*$, est noté $SL_n(\mathbb{K})$; c'est le sous-groupe des matrices dont le déterminant est 1. Autre exemple, le noyau de l'application naturelle $\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}$ est le sous-groupe $n\mathbb{Z}$.

L'une des raisons majeures, au moins dans un premier temps, pour introduire les noyaux est le petit résultat suivant, dont on se sert sans cesse :

Lemme 28. *Soit $\varphi : G_1 \longrightarrow G_2$ un homomorphisme. Alors φ est injectif si et seulement si* $\ker(\varphi) = \{e_1\}$ *(où e_1 est l'élément neutre de G_1).*

Démonstration. Supposons d'abord que φ est injectif. Alors si $g \in \ker(\varphi)$, on a $\varphi(g) = e_2 = \varphi(e_1)$, d'où $g = e_1$, et $\ker(\varphi)$ est bien réduit à $\{e_1\}$.

Réciproquement, supposons que $\ker(\varphi) = \{e_1\}$, et que $\varphi(g) = \varphi(h)$, pour $g, h \in G_1$. Alors $\varphi(gh^{-1}) = \varphi(g)\varphi(h)^{-1} = e_2$, d'où $gh^{-1} \in \ker(\varphi)$ et par suite $gh^{-1} = e_1$. On en tire bien $g = h$, et on conclut que φ est injective. □

§1.2 Anneaux

Premières définitions

Définition 29. Un *anneau* est un ensemble A muni de deux éléments distingués notés 0 et 1 (ou 0_A et 1_A quand on veut vraiment insister), avec $0 \neq 1$, et de deux opérations + et $\cdot$, ayant les propriétés suivantes :

- l'ensemble A avec l'opération + et l'élément neutre 0 est un groupe abélien (en « notation additive »),
- la multiplication $a, b \mapsto a \cdot b$ est associative, et admet 1 pour élément neutre,
- pour tous $a, b, c \in A$ on a les lois, dites de « distributivité de $\cdot$ par rapport à + », données ci-dessous :

$$a \cdot (b + c) = a \cdot b + a \cdot c, \qquad (a + b) \cdot c = a \cdot c + b \cdot c\,.$$

Enfin, on dit que A est un *anneau commutatif* lorsque c'est un anneau comme ci-dessus, et que l'on a en plus pour tous $a, b \in A$ la relation $a \cdot b = b \cdot a$. (Rappelons que l'on a $a + b = b + a$ par définition.)

Remarque. Avant même de donner des exemples, quelques remarques sur les notations et la terminologie. On va rarement écrire $a \cdot b$, et on va céder à la paresse en écrivant ab (c'est standard). Les lettres utilisées pour les anneaux seront souvent A, B, C, mais aussi (lorsque l'alphabet est « déjà pris » à cet endroit, par exemple parce qu'on a déjà des matrices A, B, C etc) les lettres R, S, T (à cause de l'anglais *ring* pour anneau).

Ajoutons que dans certains livres, un anneau tel que nous le définissons est appelé « anneau unitaire » (parce que nous supposons l'existence de l'élément 1).

Exemple 30. $\mathbb{Z}$ et $\mathbb{Z}/n\mathbb{Z}$ sont des anneaux commutatifs. Pour le second, les éléments neutres sont $\overline{0}$ et $\overline{1}$.

Exemple 31. Si A est un anneau et $n \geq 1$ est un entier, alors l'ensemble $M_n(A)$ des matrices carrées, de taille $n \times n$, à coefficients dans A, est également un anneau. Ici $0_{M_n(A)}$ est la matrice nulle (dont tous les coefficients sont nuls), et

$1_{M_n(A)}$ est la matrice identité, que l'on écrira plutot I_n ou même I. Pour $n \geq 2$, on montre que $M_n(A)$ n'est jamais commutatif, même si A est commutatif.

Exemple 32. Si A est un anneau, alors vous savez définir $A[X]$, l'ensemble des *polynômes* en X à coefficients dans A ; on montre que $A[X]$ est un anneau, pour les opérations que vous connaissez. Si A est commutatif, alors $A[X]$ est également commutatif. (A vrai dire, si A n'est pas commutatif, il est assez rare que l'on parle de $A[X]$, d'ailleurs si on le fait il faut prendre des précautions, comme préciser que X commute avec les éléments de A... dans ce cours, il n'y aura pas de tel exemple.)

Définition 33. Soit A un anneau. On dit que A est un *corps* lorsque, pour tout $a \in A$, il existe un élément $a^{-1} \in A$ tel que $aa^{-1} = a^{-1}a = 1$. L'élément a^{-1} est alors unique (exo), et on l'appelle l'inverse de a.

Exemple 34. $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des corps commutatifs (c'est-à-dire que ce sont des corps, et des anneaux commutatifs). Dans les exercices on donnera des exemples de corps non-commutatifs, ce qui n'est pas évident. Notons que $M_n(\mathbb{K})$ n'est jamais un corps pour $n \geq 2$, même si $\mathbb{K}$ est un corps commutatif. Enfin, vous devez savoir que $\mathbb{Z}/p\mathbb{Z}$ est un corps si et seulement si p est un nombre premier.

Définition 35. Si a et b sont des éléments de l'anneau A, alors on dit que a *divise* b si $b = ax$ pour un $x \in A$; on dit aussi que b est un *multiple de* a. On note parfois $a|b$. Nous en parlerons surtout dans le chapitre 3. Pour l'instant, on va juste noter que, si A est un corps et si $a \neq 0$, alors l'élément a divise tout b, puisque $b = a(a^{-1}b)$. Ajoutons que cette notion n'est presque pas employée dans les anneaux non-commutatifs.

Définition 36. Soit A un anneau, et $B \subset A$. On dit que B est un *sous-anneau* de A lorsque c'est un sous-groupe de A (pour +), qui vérifie également que $1_A \in B$, et que pour $x, y \in B$, on a $xy \in B$.

Lorsque de plus A et B sont tous les deux des corps, on dit (sans surprise) que B est un *sous-corps* de A, ou alors (c'est plus surprenant peut-être) que A est une *extension de corps* de B.

Exemple 37. $\mathbb{Q}$ est un sous-corps de $\mathbb{R}$, ou de $\mathbb{C}$; et $\mathbb{Z}$ est un sous-anneau de $\mathbb{Q}$.

Homomorphismes

Définition 38. Soient A et B deux anneaux. Un *homomorphisme d'anneaux* entre A et B est une fonction $f : A \longrightarrow B$ telle que :

- f est un homomorphisme de groupes (abéliens) entre A et B munis de l'addition ; en d'autres termes, pour $x, y \in A$ on a $f(x+y) = f(x)+f(y)$, et on en déduit $f(0_A) = 0_B$.
- pour $x, y \in A$ on a $f(xy) = f(x)f(y)$,
- $f(1_A) = 1_B$.

Le *noyau* de f est alors par définition

$$\ker(f) = \{a \in A \mid f(a) = 0_B\}.$$

Enfin, on dit sans surprise que f est un *isomorphisme d'anneaux* lorsque c'est un homomorphisme d'anneaux et une bijection (on montre alors que f^{-1} est aussi un homomorphisme).

Exemple 39. Soit A un anneau et $B = M_n(A)$. On définit $f : A \longrightarrow B$ par $f(a) = aI$, où I est la matrice identité $n \times n$. Vous vérifierez alors que f est un homomorphisme d'anneaux.

Autre exemple, la projection naturelle $\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}$ est un homomorphisme d'anneaux : c'est une façon de rappeler les identités familières $\overline{a+b} = \overline{a} + \overline{b}$ et $\overline{ab} = \overline{a}\,\overline{b}$.

Notons qu'un homomorphisme d'anneaux $f : A \longrightarrow B$ est en particulier un homomorphisme de groupes, si l'on voit A et B comme des groupes abéliens pour +, et la définition du noyau de f est la même que celle que nous avions donnée pour les groupes. Nous avons donc, en appliquant le lemme 28 :

Lemme 40. *Soit $f : A \longrightarrow B$ un homomorphisme d'anneaux. Alors f est injectif si et seulement si* $\ker(f) = \{0\}$. □

Il n'y a rien à démontrer !

Voici une application qui vous servira énormément au second semestre :

Lemme 41. *Soit $f : \mathbb{K} \longrightarrow A$ un homomorphisme d'anneaux, où l'on suppose que $\mathbb{K}$ est un corps (et A un anneau quelconque). Alors f est injectif.*

Démonstration. Soit $x \in \ker(f)$. Si on avait $x \neq 0$, alors on pourrait prendre l'inverse x^{-1} et faire le calcul suivant :

$$f(xx^{-1}) = f(1) = 1 = f(x)f(x^{-1}) = 0 \cdot f(x^{-1}) = 0\,.$$

C'est bien sûr absurde. (On a utilisé ici le fait que, dans un anneau, $0 \cdot a = 0$ pour tout a.) Cette contradiction montre que $x = 0$, donc que $\ker(f) = \{0\}$, et ainsi f est injectif d'après le lemme. □

Algèbres

Définition 42. Soit $\mathbb{K}$ un corps commutatif, et soit A un anneau. On dit que A est *une algèbre sur* $\mathbb{K}$ lorsqu'il existe un homomorphisme $\iota : \mathbb{K} \longrightarrow A$ ayant la propriété que, pour tout $\lambda \in \mathbb{K}$ et tout $a \in A$, on a $\iota(\lambda) \cdot a = a \cdot \iota(\lambda)$.

Exemple 43. L'anneau $A = M_n(\mathbb{K})$ des matrices carrées à coefficients dans $\mathbb{K}$ est une algèbre sur $\mathbb{K}$, avec $\iota(\lambda) = \lambda I$, en écrivant I pour la matrice identité. Autre exemple, l'anneau $A = \mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$ est une algèbre sur $\mathbb{K}$, avec $\iota(\lambda) = \lambda$ (le polynôme constant égal à λ).

Ici nous allons faire un abus de notation assez important. En effet, si A est une algèbre sur $\mathbb{K}$, alors l'homomorphisme ι est injectif, par le dernier lemme. Il réalise donc un isomorphisme entre $\mathbb{K}$ et $\iota(\mathbb{K})$. L'abus de notation consiste alors à ne plus parler de ι, et à considérer $\mathbb{K}$ comme un sous-anneau de A. Lorsque A est une algèbre sur $\mathbb{K}$, et que $\lambda \in \mathbb{K}$, $a \in A$, on s'autorise donc à écrire $\lambda \cdot a$ au lieu de $\iota(\lambda) \cdot a$.

Dans les exemples ci-dessus, ceci revient à identifier $\mathbb{K}$ avec les matrices de la forme λI, pour $\lambda \in \mathbb{K}$, dans le premier cas, ou alors avec l'ensemble des polynômes constants dans le deuxième cas. En pratique, à chaque fois que nous énoncerons « A est une algèbre sur $\mathbb{K}$ », l'homomorphisme ι sera évident, ou plutôt, il y aura un sous-anneau de A isomorphe à $\mathbb{K}$ qui sera le candidat évident.

Définition 44. Soient A et B deux algèbres sur $\mathbb{K}$. On dit que $f : A \longrightarrow B$ est un *homomorphisme d'algèbres sur* $\mathbb{K}$ lorsque c'est un homomorphisme d'anneaux qui vérifie en plus, pour $\lambda \in \mathbb{K}$ et $a \in A$, la relation $f(\lambda \cdot a) = \lambda \cdot f(a)$.

(Cette définition utilise l'abus de notation juste mentionné.)

Nous allons voir un exemple fondamental d'homomorphisme d'algèbres, impliquant $\mathbb{K}[X]$. Nous avons besoin d'une notation, que vous connaissez : soit A une algèbre sur $\mathbb{K}$, soit $P = \lambda_0 + \lambda_1 X + \cdots + \lambda_n X^n \in \mathbb{K}[X]$, et soit $a \in A$; alors $\lambda_0 + \lambda_1 a + \cdots + \lambda_n a^n$ est noté $P(a)$. (On dit qu'on évalue P en a.)

Nous pouvons alors énoncer :

Proposition 45. *Soit $\mathbb{K}$ un corps commutatif, et A une algèbre sur A. Alors tout homomorphisme d'algèbres $\mathbb{K}[X] \longrightarrow A$ est de la forme $P \mapsto P(a)$ pour un unique $a \in A$. (On parle de l'évaluation en a.)*

Démonstration. Si $a \in A$ nous est donné, le fait que $P \mapsto P(a)$ est un homomorphisme d'algèbres est (censé être) évident d'après la définition des opérations sur $\mathbb{K}[X]$. On affirme simplement que $(P + Q)(a) = P(a) + Q(a)$ et $(PQ)(a) = P(a)Q(a)$. Par ailleurs, pour $P = X$, on a $P(a) = a$ donc l'élément a est déterminé par l'homomorphisme ; il est donc bien unique.

Ce qui nous intéresse ici, c'est de montrer qu'il n'y a pas d'autres homomorphismes. Soit $f : \mathbb{K}[X] \longrightarrow A$ un homomorphisme. On a

$$f(\lambda_0 + \lambda_1 X + \cdots + \lambda_n X^n) = \lambda_0 + \lambda_1 f(X) + \cdots + \lambda_n f(X)^n,$$

ce qui montre bien, si on pose $a = f(X)$, que f est simplement l'évaluation en a. □

Exemple 46. Prenons $\mathbb{K} = \mathbb{R}$ et $A = M_2(\mathbb{R})$. Donnons-nous la matrice

$$M = \begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix}.$$

Nous avons donc un homomorphisme d'évaluation en M, noté $P \mapsto P(M)$, de $\mathbb{R}[X]$ vers $M_2(\mathbb{R})$. Si l'on choisit par exemple

$$P = 2X^2 - 7X + 4,$$

voyons ce que donne $P(M)$. On a

$$M^2 = \begin{pmatrix} -2 & 6 \\ -3 & 7 \end{pmatrix}.$$

Par ailleurs, il faut réfléchir un peu pour interpréter le « +4 » : ici $M_2(\mathbb{R})$ est vu comme une algèbre sur $\mathbb{R}$ de la façon usuelle (comme ci-dessus), c'est-à-dire qu'on identifie $\mathbb{R}$ avec l'ensemble des matrices de la forme λI, où $\lambda \in \mathbb{R}$ et I est la matrice identité 2×2.

Finalement

$$P(M) = 2M^2 - 7M + 4I = 2\begin{pmatrix} -2 & 6 \\ -3 & 7 \end{pmatrix} - 7\begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix} + 4\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix}.$$

Chapitre 2

Modules

§2.1 Premières définitions

Définition 47. Soit A un anneau et V un groupe abélien. Une *structure de A-module* sur V est une application

$$A \times V \longrightarrow V\,,$$

notée $(a, v) \mapsto a \cdot v$, avec les propriétés suivantes :

- l'opération est *bilinéaire*, c'est-à-dire que l'on a

  $$(a+b)\cdot v = a\cdot v + b\cdot v\,, \qquad a\cdot(v+w) = a\cdot v + a\cdot w\,,$$

  pour $a, b \in A$ et $v, w \in V$;
- l'opération est associative, dans le sens où

  $$a\cdot(b\cdot v) = (ab)\cdot v$$

  pour $a, b \in A$ et $v \in V$, en écrivant bien sûr ab pour la multiplication dans A ;
- $1 \cdot v = v$ pour tout $v \in V$.

La première chose à remarquer, et c'est capital, c'est que dans la situation où A est un corps, noté disons $\mathbb{K}$, dire que V possède une structure de $\mathbb{K}$-module revient exactement à dire que V est un espace vectoriel sur $\mathbb{K}$. C'est donc quelque chose que vous connaissez bien ! Le but de ce chapitre, et plus généralement du cours d'algèbre de ce semestre, est de voir si l'on peut « faire

de l'algèbre linéaire » avec d'autres anneaux, qui ne sont pas des corps. Nous verrons que certaines notions se généralisent bien, mais dans l'ensemble le cas des corps est très particulier...

Voyons donc d'autres exemples. Pour cela, il sera utile de reformuler un peu la définition ci-dessus, car elle n'est pas toujours la plus pratique, selon les situations. Rappelons que, au cours des exercices sur le chapitre précédent, nous avons vu que pour tout groupe abélien V, l'ensemble $\mathrm{End}(V)$ des endomorphismes de V est naturellement muni d'une structure d'anneau, la multiplication étant la composition $\circ$, et l'élément neutre de cette multiplication étant l'identité I.

Lemme 48. *Soit V un groupe abélien et A un anneau. Se donner une structure de A-module sur V revient exactement à se donner un homomorphisme d'anneaux $\rho : A \longrightarrow \mathrm{End}(V)$.*

Démonstration. Supposons que V est un A-module au sens précédent. Pour chaque $a \in A$, notons

$$\begin{aligned} \rho(a) : V &\longrightarrow V \\ v &\mapsto a \cdot v \,. \end{aligned}$$

Ainsi $\rho(a)$ est un élément de $\mathrm{End}(V)$, et $a \mapsto \rho(a)$ est un homomorphisme d'anneaux : ces deux affirmations découlent des propriétés ci-dessus des A-modules (vérifiez-le).

Réciproquement, supposons donné $\rho : A \longrightarrow \mathrm{End}(V)$. Pour éviter les lourdeurs, nous écrirons $a \mapsto \rho_a$ (et non pas $\rho(a)$). Il suffit alors de poser, pour $a \in A$ et $v \in V$:

$$a \cdot v := \rho_a(v) \,.$$

Vous vérifierez (c'est le même calcul que ci-dessus, mais à l'envers) que $(a, v) \mapsto a \cdot v$ est bien une structure de A-module.

Enfin, il est immédiat que ces deux constructions sont inverses l'une de l'autre. □

Exemple 49. Qu'est-ce qu'un $\mathbb{Z}$-module ? Il faut prendre un groupe abélien V, et trouver un homomorphisme $\mathbb{Z} \longrightarrow \mathrm{End}(V)$. Or, nous avons vu ça dans

les exercices du chapitre précédent, pour tout anneau A il existe un *unique* homomorphisme $\mathbb{Z} \longrightarrow A$; ici pour $A = \mathrm{End}(V)$, il s'agit de $n \mapsto nI$. Donc V est automatiquement un $\mathbb{Z}$-module, de manière unique. En bref, *un $\mathbb{Z}$-module n'est rien d'autre qu'un groupe abélien.*

Encore quelques préliminaires avant un autre exemple important.

Définition 50. Soient V et W des A-modules. Une application $f : V \longrightarrow W$ est appelée *homomorphisme de A-modules* lorsque c'est un homomorphisme de groupes abéliens et que $f(a \cdot v) = a \cdot f(v)$ pour $a \in A$ et $v \in V$. On dit que f est *A-linéaire*. Lorsque $V = W$, on dit que f est un *endomorphisme* du A-module V. Enfin, on dit que f est un *isomorphisme de A-modules* lorsque c'est un homomorphisme et une bijection.

Comme prévu, dans le cas où A est un corps, vous retrouvez une notion familière.

Lemme 51. *Soit V un A-module, et soit $\mathrm{End}_A(V)$ l'ensemble de ses endomorphismes de A-module. Alors $\mathrm{End}_A(V)$ est un sous-anneau de $\mathrm{End}(V)$. Lorsque $A = \mathbb{K}$ est un corps commutatif, $\mathrm{End}_{\mathbb{K}}(V)$ est même une algèbre sur $\mathbb{K}$.*

Démonstration. C'est très simple : il faut prendre $f, g \in \mathrm{End}_A(V)$ et écrire que

$$f \circ g(a \cdot v) = f(g(a \cdot v)) = f(a \cdot g(v)) = a \cdot f(g(v)) = a \cdot f \circ g(v)\,,$$

ce qui montre bien que $f \circ g \in \mathrm{End}_A(V)$.

Supposons maintenant que $A = \mathbb{K}$ est un corps commutatif (on parle donc d'espaces vectoriels ici). Écrivons I pour l'identité de V, et pour $\lambda \in \mathbb{K}$ écrivons sans surprise λI pour l'application $v \mapsto \lambda \cdot v$. L'ensemble des λI avec $\lambda \in \mathbb{K}$ est un anneau, et même un sous-anneau de $\mathrm{End}_{\mathbb{K}}(V)$ (car $\mathbb{K}$ est commutatif !) que l'on peut identifier avec $\mathbb{K}$. Ceci donne bien à $\mathrm{End}_{\mathbb{K}}(V)$ une structure d'algèbre sur $\mathbb{K}$: en effet il faut vérifier que $\lambda I \circ f = f \circ \lambda I$ pour $f \in \mathrm{End}_{\mathbb{K}}(V)$, mais cette condition revient exactement à dire que f est $\mathbb{K}$-linéaire. □

Tournons-nous vers le cas des algèbres, et commençons par une remarque. Lorsque B est un sous-anneau de A, tout A-module peut être considéré comme un B-module, si l'on veut : par exemple un espace vectoriel sur $\mathbb{C}$ est aussi un

espace vectoriel sur $\mathbb{R}$. Prenons alors un corps commutatif $\mathbb{K}$ et une algèbre sur $\mathbb{K}$, notée A. Puisque A possède un sous-anneau que l'on identifie à $\mathbb{K}$, la remarque précédente montre que tout A-module est en particulier un espace vectoriel sur $\mathbb{K}$. On peut alors énoncer le résultat suivant, qui est la variante pour les algèbres du lemme 48.

Lemme 52. *Soit A une algèbre sur $\mathbb{K}$, et soit V un groupe abélien. Se donner une structure de A-module sur V revient exactement à se donner une structure de $\mathbb{K}$-espace vectoriel sur V ainsi qu'un homomorphisme d'algèbres $\rho : A \longrightarrow \mathrm{End}_{\mathbb{K}}(V)$.*

Démonstration. On vous le laisse à titre d'exercice. C'est une variante de la démonstration du lemme 48. Attention à une chose : la définition de $\mathrm{End}_{\mathbb{K}}(V)$ dépend bel et bien de la structure de $\mathbb{K}$-espace vectoriel choisie. □

Exemple 53. Qu'est-ce qu'un $\mathbb{K}[X]$-module ? D'après le lemme, il s'agit d'un $\mathbb{K}$-espace vectoriel V muni d'un homomorphisme $\rho : \mathbb{K}[X] \longrightarrow \mathrm{End}_{\mathbb{K}}(V)$. Mais une proposition du chapitre précédent nous dit qu'un tel homomorphisme d'algèbre est de la forme $P \mapsto P(f)$, où $f \in \mathrm{End}_{\mathbb{K}}(A)$, et qu'il suffit de nous donner f. En bref, *un $\mathbb{K}[X]$-module est une paire (V, f) où V est un espace vectoriel sur $\mathbb{K}$ et $f : V \longrightarrow V$ est un $\mathbb{K}$-endomorphisme.* Pour être tout-à-fait clair, pour $v \in V$ on a

$$X \cdot v = f(v)\,.$$

Pour $P \in \mathbb{K}[X]$ quelconque, on a alors

$$P \cdot v = P(f)(v)\,;$$

ces parenthèses sont bien pénibles, mais $P(f)$ est un endomorphisme de V, on peut donc l'appliquer à v pour obtenir $P(f)(v)$. Voir l'exemple suivant.

Exemple 54. Soyons très concrets, et définissons un $\mathbb{R}[X]$-module. On prend $V = \mathbb{R}^2$ (dans tout ce cours, les éléments de $\mathbb{K}^n$ sont vus comme des matrices-colonnes, au fait). Il nous faut un endomorphisme $f : V \longrightarrow V$. Comme on le sait bien, il doit être de la forme $f(v) = Fv$, où F est la matrice de f dans la base canonique ; c'est une matrice 2×2, et ici Fv désigne bien le produit matriciel. (Note : la lettre F est plutôt rare pour une matrice, mais dans ce cours on

va essayer, dans la mesure du possible, d'appeler F la matrice de f, puis G la matrice de g etc.) Pour continuer à être concret, on peut prendre par exemple

$$F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Si on prend un polynôme, disons $P = X^3 - 4$, alors la matrice de l'endomorphisme $P(f)$ est tout simplement $P(F) = F^3 - 4I$. De sorte que si l'on prend un vecteur v, on a

$$P \cdot v = (X^3 - 4) \cdot v = (F^3 - 4I)v \qquad (= P(f)(v)).$$

Ici

$$F^3 - 4I = \begin{pmatrix} -3 & 3 \\ 0 & -3 \end{pmatrix},$$

donc si on prend disons $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, alors

$$P \cdot v = \begin{pmatrix} 3 \\ -3 \end{pmatrix}.$$

Maintenant que nous pouvons penser aux $\mathbb{K}[X]$-modules comme à des paires (V, f), nous pouvons « traduire » la notion d'homomorphisme :

Lemme 55. *Soient (V, f) et (W, g) deux $\mathbb{K}[X]$-modules, et soit $\varphi : V \longrightarrow W$. Alors φ est un homomorphisme de $\mathbb{K}[X]$-modules si et seulement si φ est $\mathbb{K}$-linéaire et vérifie $\varphi \circ f = g \circ \varphi$.*

La situation est comme sur le diagramme suivant :

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ W & \xrightarrow{g} & W \end{array}$$

Démonstration. Si φ est $\mathbb{K}[X]$-linéaire, alors elle est certainement $\mathbb{K}$-linéaire ; et de plus, on doit avoir pour $v \in V$ la relation

$$\varphi(X \cdot v) = X \cdot \varphi(v).$$

Mais ici on a $X \cdot v = f(v)$, et pour tout $w \in W$ on a $X \cdot w = g(w)$, donc finalement

$$\varphi(f(v)) = g(\varphi(v)),$$

comme on le souhaitait.

Pour la réciproque, on suppose que φ est $\mathbb{K}$-linéaire et vérifie $\varphi \circ f = g \circ \varphi$, et on doit montrer pour tout polynôme $P \in \mathbb{K}[X]$ que $\varphi(P \cdot v) = P \cdot \varphi(v)$ pour tout v. C'est certainement vrai pour $P = X$, puisque la condition $\varphi(X \cdot v) = X \cdot \varphi(v)$ n'est que la traduction de $\varphi \circ f = g \circ \varphi$. Or l'ensemble

$$R = \{P \in \mathbb{K}[X] \mid \varphi(P \cdot v) = P \cdot \varphi(v)\}$$

est, à l'évidence, un sous-anneau de $\mathbb{K}[X]$. Puisque $X \in R$, on en déduit que $R = \mathbb{K}[X]$ en entier. □

Exemple 56. Pour $W = V$ et $g = f$, la condition est que $\varphi \circ f = f \circ \varphi$, c'est-à-dire que φ et f commutent. Reprenons l'exemple 54. Si Φ est la matrice de $\varphi : V \longrightarrow V$, alors la condition pour que φ soit un endomorphisme de (V, f) peut s'écrire $\Phi F = F\Phi$. Si

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

alors en écrivant séparément ΦF et $F\Phi$ on voit que la condition équivaut à $c = 0$ et $a = d$. En d'autres termes, nous avons une description de l'anneau $\mathrm{End}_{\mathbb{R}[X]}(V)$ des endomorphismes de (V, f) comme $\mathbb{R}[X]$-module :

$$\mathrm{End}_{\mathbb{R}[X]}(V) = \left\{ \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \text{ avec } a, b \in \mathbb{R} \right\}.$$

C'est une algèbre sur $\mathbb{R}$, et comme espace vectoriel réel, elle est de dimension 2.

Nous pouvons maintenant annoncer plus clairement nos intentions dans ce cours. Nous allons parler de modules tout le long du semestre. Pour chaque résultat ou concept, ou presque, concernant les A-modules :

- lorsque A est un corps, on retrouvera quelque chose de connu d'algèbre linéaire, et en général on pourra réduire l'énoncé à quelque chose de beaucoup plus simple ;

- lorsque $A = \mathbb{Z}$, les choses seront immédiatement plus subtiles : on aura besoin des « vrais » énoncés généralisés, car si on s'attendait à ce que l'algèbre linéaire fonctionne « de la même manière » sur $\mathbb{Z}$, on aurait bien tort, les contre-exemples étant abondants ;
- lorsque $A = \mathbb{K}[X]$, les questions naturelles sur les modules se traduisent en questions que vous avez déjà un peu étudiées en cours de « réduction des endomorphismes » ; comme vous le savez, certaines de ces questions sont sophistiquées.

Nous étudierons souvent les A-modules sans hypothèse sur A, mais ce sont les 3 situations ci-dessus que nous allons systématiquement examiner dans les exemples. Puis, arrivera un chapitre où l'on montrera que $\mathbb{Z}$ et $\mathbb{K}[X]$ ne sont pas si différents (ce sont des « anneaux euclidiens »), et que leurs modules ne sont pas si compliqués, après tout !

La dernière partie du cours va étudier les « G-modules », où G est un groupe. Ceux-ci sont très bien compris, et avec la théorie des « caractères » on peut même rendre les choses très concrètes.

§2.2 Sous-modules

Dans cette partie, A est un anneau quelconque, et $\mathbb{K}$ est un corps commutatif.

Définition 57. Soit V un A-module, et $U \subset V$. On dit que U est un *sous-A-module de V* (ou sous-module de V pour faire court) lorsque c'est un sous-groupe de V, et que pour $u \in U$ et $a \in A$ on a $a \cdot u \in U$. (En d'autres termes, U est « stable par combinaisons linéaires à coefficients dans A ».)

Dans ce cas U est lui-même un A-module.

Exemple 58. Pour $A = \mathbb{K}$, on retrouve la notion de sous-espace vectoriel. Pour $A = \mathbb{Z}$, on retrouve la notion de sous-groupe (abélien). Lorsque $A = \mathbb{K}[X]$, et que (V, f) est donc un $\mathbb{K}[X]$-module, dire que $U \subset V$ est un sous-$\mathbb{K}[X]$-module de V signifie exactement que $f(U) \subset U$, autrement dit que U est stable par f. En écrivant $f|_U$ pour la restriction de f à U, on a bien un $\mathbb{K}[X]$-module $(U, f|_U)$.

Définition 59. Soient U_1 et U_2 deux sous-modules de V. La *somme* $U_1 + U_2$ est par définition

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}\,.$$

C'est un sous-module de V. On dit que V est *la somme directe* de U_1 et U_2, et on écrit $V = U_1 \oplus U_2$, lorsque $V = U_1 + U_2$ et $U_1 \cap U_2 = \{0\}$.

Avant même de donner des exemples, nous pouvons rappeler le fait suivant, qui vous est familier dans le cas $A = \mathbb{K}$ et qui n'est pas plus compliqué en général :

Lemme 60. *Soient U_1 et U_2 deux sous-modules de V. Alors $V = U_1 \oplus U_2$ est équivalent au fait que tout $v \in V$ peut s'écrire $v = u_1 + u_2$ avec $u_i \in U_i$, et ceci de manière unique.*

Démonstration. Exercice. Comme dans le cas $A = \mathbb{K}$ que vous connaissez. □

Exemple 61. Pour $A = \mathbb{K}$, on retrouve la notion de somme directe que vous connaissez depuis longtemps. Regardons un peu $A = \mathbb{Z}$. Prenons $V = \mathbb{Z}$, qui est bien un groupe abélien. Tout sous-module (= sous-groupe, ici) de $\mathbb{Z}$ est de la forme $n\mathbb{Z}$ pour un entier $n \geq 0$. Si $U_1 = n\mathbb{Z}$ et $U_2 = m\mathbb{Z}$, avec n et m tous les deux > 0, alors $U_1 \cap U_2$ n'est jamais réduit à $\{0\}$, puisqu'il contient toujours $nm \neq 0$ par exemple. Donc on ne peut pas écrire V comme une somme directe, à part si U_1 ou U_2 est nul. (Dans les exercices nous étudierons $U_1 + U_2$ et $U_1 \cap U_2$.)

Enfin, soit (V, f) un $\mathbb{K}[X]$-module, avec V de dimension finie sur $\mathbb{K}$. Supposons que $V = U_1 \oplus U_2$, où U_i est un sous-$\mathbb{K}[X]$-module. Prenons une base de U_1, disons $e_1, \ldots, e_d$, et une base de U_2, disons $\varepsilon_1, \ldots, \varepsilon_r$. Alors la réunion $e_1, \ldots, e_d, \varepsilon_1, \ldots, \varepsilon_r$ est une base de V. Dans cette base, la matrice de f est de la forme

$$\begin{pmatrix} F_1 & 0 \\ 0 & F_2 \end{pmatrix}$$

où F_1 est une matrice $d \times d$ et F_2 est une matrice $r \times r$; en fait F_i est la matrice de $f|_{U_i}$. Autrement dit *il existe une base dans laquelle la matrice de f est diagonale par blocks*. Réciproquement, s'il existe une telle base, il est clair que V peut s'exprimer comme une somme directe.

Définition 62. Un A-module V est dit *indécomposable* lorsqu'on ne peut pas trouver de sous-modules U_1 et U_2, tous les deux non-nuls, tels que $V = U_1 \oplus U_2$.

L'exemple précédent montre que $\mathbb{Z}$, comme $\mathbb{Z}$-module, est indécomposable. Voyons un autre exemple.

Exemple 63. Revenons à la situation de l'exemple 54, et montrons que V est indécomposable. Supposons donc que $V = U_1 \oplus U_2$. Comme espace vectoriel sur $\mathbb{K}$, nous savons que V est de dimension 2 ; si les deux U_i sont non-nuls, on doit donc avoir $\dim U_1 = \dim U_2 = 1$. Mais alors, tout vecteur non-nul $u_i \in U_i$ est un *vecteur propre* de f, puisque $f(u_i) \in U_i = \mathcal{V}ect(u_i)$. On aurait donc une base u_1, u_2 de vecteurs propres de f, ou autrement dit, f serait diagonalisable. Or, rappelons que la matrice de f est

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

son polynôme caractèristique est $(X-1)^2$, la seule valeur propre est 1, et l'espace propre $\ker(f - I)$ est de dimension 1, donc f n'est pas diagonalisable. (Ou encore : avec une seule valeur propre, si f était diagonalisable, elle serait déjà diagonale, comme on le sait bien.) Cette contradiction montre que V est indécomposable.

Dans le cas $A = \mathbb{K}$, c'est-à-dire dans le cas des espaces vectoriels, quels sont les modules indécomposables ? Il y en a très peu, en conséquence de la proposition suivante :

Proposition 64. *Soit V un espace vectoriel [de dimension finie pour simplifier], et soit U un sous-espace de V. Alors il existe un sous-espace U' tel que $V = U \oplus U'$.*

Ici, et dans le reste de cette partie, nous ferons des hypothèses de dimension finie, qui ne sont pas nécessaires : le résultat est vrai en toute généralité, mais il faut l'axiome du choix et diverses choses que nous préférons éviter... On va essayer d'en rester à des résultats que vous avez vus en L1.

Démonstration. Soit $e_1, \ldots, e_d$ une base de U – il en existe, d'après votre cours d'algèbre linéaire de L1. On utilise le théorème de la base incomplète, qui

nous affirme que l'on peut trouver des vecteurs $e_{d+1}, \ldots, e_n$ tels que $e_1, \ldots, e_n$ est une base de V. Si l'on pose $U' = \mathcal{V}ect(e_{d+1}, \ldots, e_n)$, alors il est clair que $V = U \oplus U'$. □

Dans cette situation, dès lors que V possède un sous-espace U qui est non-nul, et tel que $U \neq V$, alors $V = U \oplus U'$ montre que V n'est pas indécomposable. On sent qu'il ne va pas y avoir beaucoup de modules indécomposables !

Le vocabulaire suivant va être utile :

Définition 65. Un module V non-nul est dit *simple* lorsque, pour tout sous-module $U \subset V$, on a ou bien $U = V$ ou bien $U = \{0\}$.

On a donc toujours :

Lemme 66. *Si V est un A-module simple, alors V est indécomposable.*

Démonstration. En effet, si on a $V = U_1 \oplus U_2$, alors par simplicité de V, on doit avoir ou bien $U_1 = V$ (et donc $U_2 = \{0\}$) ou bien $U_1 = \{0\}$ (et $U_2 = V$) ; on voit que U_1 et U_2 ne peuvent pas être tous les deux non-nuls, et donc que V est bel et bien indécomposable. □

Pour $A = \mathbb{K}$, la situation est complètement sous-contrôle :

Proposition 67. *Soit V un espace vectoriel sur $\mathbb{K}$ non-nul [de dimension finie pour simplifier]. Alors les trois propriétés ci-dessous sont équivalentes :*

1. *V est simple,*
2. *V est indécomposable,*
3. *V est de dimension* 1.

Démonstration. D'après le dernier lemme, on a toujours (1) $\implies$ (2). Supposons (2), et soit $v \in V$, $v \neq 0$ et $U = \mathcal{V}ect(v)$. D'après la proposition, on peut trouver U' tel que $V = U \oplus U'$. Mais comme V est indécomposable, ceci n'est possible que si $U' = \{0\}$. On en déduit que $V = U = \mathcal{V}ect(v)$, et en particulier la dimension de V est (1).

Enfin, supposons (3). Un sous-espace U de V doit être de dimension inférieure à 1, donc soit $\dim U = 0$ et $U = \{0\}$, soit $\dim U = 1$ et $U = V$. On a bien montré (1). □

On comprend bien pourquoi, dans vos cours d'algèbre linéaire, on ne vous a pas embêtés avec les notions de « module indécomposable » et « module simple » ! Mais sur un anneau quelconque, les choses sont plus délicates.

Exemple 68. Pour $A = \mathbb{Z}$, nous avons vu ci-dessus que $V = \mathbb{Z}$ est indécomposable. Mais il n'est pas simple : pour tout $n > 0$, le sous-groupe $U = n\mathbb{Z}$ est non-nul, et $U \neq V$ (il y a donc une infinité de sous-modules qui contredisent la simplicité de V).

Voyons des $\mathbb{Z}$-modules simples. Si p est un entier, alors les sous-groupes de $V = \mathbb{Z}/p\mathbb{Z}$ sont en bijection avec les diviseurs de p ; plus précisément, si $p = dk$, alors $U = \{v \in V \mid dv = 0\}$ est l'unique sous-groupe de V d'ordre d. Par conséquent, si p est un nombre premier, alors $\mathbb{Z}/p\mathbb{Z}$ est un module simple (on dit aussi un groupe simple, ici).

Il y a donc une infinité de modules simples (un pour chaque nombre premier), qui ne sont pas isomorphes les uns aux autres (alors que dans le cas des corps, les modules simples, ie les modules de dimension 1, sont bien sûr tous isomorphes les uns aux autres). On verra plus loin qu'il n'y a pas d'autres $\mathbb{Z}$-modules simples.

Exemple 69. Voyons maintenant le cas de $A = \mathbb{K}[X]$. Retournons encore une fois à l'exemple 54. Nous venons de voir que ce module est indécomposable. Par contre, il n'est pas simple : si e_1, e_2 est la base canonique de V comme $\mathbb{R}$-espace vectoriel, alors $U = \mathcal{V}ect(e_1)$ est stable par f, visiblement, donc U est un sous-$\mathbb{K}[X]$-module de V.

Pour produire un exemple de $\mathbb{K}[X]$-module simple, il suffit de prendre V de dimension 1, avec f donné par une matrice 1×1 (bref un scalaire λ). En effet, un sous-module U d'un tel V doit être en particulier un sous-espace vectoriel, mais puisque V est alors simple comme $\mathbb{K}$-module (par la proposition), on a certainement $U = V$ ou $U = \{0\}$. En faisant varier λ, on obtient une collection de modules simples qui ne sont pas isomorphes les uns aux autres.

Plus loin dans ce cours, nous donnerons une classification des $\mathbb{Z}$-modules et des $\mathbb{K}[X]$-modules (avec un seul théorème !), et nous pourrons alors décrire complètement les indécomposables et les simples.

Terminons cette partie avec la définition du produit de deux modules :

Définition 70. Soient V et W deux modules sur A. Alors leur produit cartésien $V \times W$ est vu comme un A-module avec la structure

$$a \cdot (v, w) = (a \cdot v, a \cdot w)$$

pour $a \in A$, $v \in V$, $w \in W$ (et avec la structure naturelle de groupe abélien sur $V \times W$). On l'appelle tout naturellement le produit de V et W.

La notation $V \times W$ va être abusée presque tout de suite (dans ce cours, on va essayer de faire attention, mais c'est vrai que c'est tentant). En effet, le sous-module

$$V \times \{0\} = \{(v, 0) \mid v \in V\}$$

peut être sans danger identifié à V, et de même on identifie $\{0\} \times W$ à W. Ayant fait ceci, on voit V et W comme des sous-modules de $V \times W$, et ils sont alors en somme directe : $V \times W = V \oplus W$. Voilà pourquoi on trouve souvent la notation $V \oplus W$ là où il serait plus juste d'écrire $V \times W$.

Ajoutons enfin qu'il existe des définitions de la somme directe $\bigoplus_{i \in I} U_i$ d'une famille quelconque de modules U_i indexés par l'ensemble I, ainsi que du produit $\prod_{i \in I} U_i$ de ces mêmes modules : nous n'en dirons rien dans ce cours, mais sachez que ces deux constructions sont bien distinctes. (Alors que pour deux modules, ou même pour un nombre fini de modules, on vient de voir que confondre produit et somme directe n'est pas dramatique.)

§2.3 Modules libres

A désigne un anneau, et $\mathbb{K}$ est un corps commutatif.

Généralités

Définition 71. Pour tout entier $n \geq 1$, on écrit A^n pour le produit de n copies de A, c'est-à-dire le produit cartésien formé des n-uplets $(a_1, \ldots, a_n)$ avec $a_i \in A$; c'est un A-module avec

$$a \cdot (a_1, \ldots, a_n) = (aa_1, \ldots, aa_n)\,.$$

On dit qu'un module V est *libre de rang n* s'il est isomorphe à A^n.

Pour $n = 1$, le module A^1 n'est autre que A, vu comme module sur lui-même (!), en utilisant la multiplication. On l'appelle parfois le *module régulier* de A. Il sera parfois utile de garder la notation A^1 pour le module régulier, quand on veut le distinguer de l'anneau A. Noter que A^n est le produit de n copies du module A^1.

Voici des concepts qui vous sont familiers :

Définition 72. Soit V un A-module, et $v_1, \ldots, v_n$ une famille d'éléments de V.

1. On dit que $v_1, \ldots, v_n$ est une *famille génératrice* lorsque l'application
$$\begin{aligned} A^n &\longrightarrow V \\ (a_1, \ldots, a_n) &\mapsto a_1 v_1 + \cdots + a_n v_n \end{aligned}$$
est surjective.
2. On dit que $v_1, \ldots, v_n$ est une *famille libre* lorsque l'application ci-dessus est injective.
3. On dit que $v_1, \ldots, v_n$ est une *base* de V lorsque c'est une famille à la fois libre et génératrice, ou en d'autres termes lorsque l'application ci-dessus est un isomorphisme.

Prenez le temps de bien vérifier que ceci correspond à la façon dont vous avez vu ces concepts dans le cadre de l'algèbre linéaire. Par exemple, pour le deuxième point, rappelez-vous bien que l'application que l'on regarde est injective $\Longleftrightarrow$ son noyau est réduit à $\{0\}$, c'est-à-dire si et seulement si la seule façon d'avoir une combinaison linéaire nulle
$$a_1 v_1 + \cdots + a_n v_n = 0$$
est de prendre tous les a_i nuls. (L'élément neutre 0 dans le module A^n est $(0, 0, \ldots, 0)$, évidemment !)

Vérifions que nous comprenons bien ce vocabulaire :

Lemme 73. *Soit V un A-module. Alors V est libre de rang n si et seulement s'il possède une base formée de n éléments.*

Démonstration. Si V possède une base avec n éléments, par définition il est isomorphe à A^n, donc libre. Pour la réciproque, dans le module A^n on peut utiliser les éléments

$$e_i = (0, \dots, 0, 1, 0, \dots, 0)$$

avec le 1 en i-ième position. La famille $e_1, \dots, e_n$ est une base de A^n, évidemment. Si $\varphi : A^n \longrightarrow V$ est un isomorphisme, alors la famille $\varphi(e_1), \dots, \varphi(e_n)$ est une base de V. □

Définition 74. Un A-module V est dit *de type fini* lorsqu'il possède une famille génératrice (finie).

Lorsque $A = \mathbb{K}$, on dit plutôt que V est de *dimension finie* plutôt que de type fini, comme vous le savez. Un résultat très fort d'algèbre linéaire est le suivant :

Proposition 75. *Tout espace vectoriel de dimension finie possède une base.* □

Là encore, la vérité est qu'on n'a pas besoin de supposer que l'espace vectoriel est de dimension finie, mais vous n'avez sans doute pas vu la version plus générale. La moralité est que *sur un corps, tous les modules sont libres.*

Comme d'habitude, c'est loin d'être le cas avec les autres anneaux. Avec $A = \mathbb{Z}$, il suffit de prendre $V = \mathbb{Z}/2\mathbb{Z}$, qui ne risque pas d'être libre, c'est-à-dire isomorphe à $\mathbb{Z}^n$ pour un certain n, puisqu'il n'est même pas infini ! Et dans la même veine, avec $\mathbb{K}[X]$ il suffit de prendre une paire (V, f) avec V de dimension finie sur $\mathbb{K}$: il n'a alors aucune chance d'être isomorphe à $\mathbb{K}[X]^n$, qui est de dimension infinie sur $\mathbb{K}$.

Idéaux

Puisque nous venons d'introduire le module régulier A^1, nous pouvons parler des *idéaux*, qui fournissent de bons exemples de modules :

Définition 76. Les sous-modules de A^1 sont appelés les *idéaux* de A (ou parfois les *idéaux à gauche*, pour être plus précis). En clair, un idéal de A est un sous-groupe $I \subset A$ ayant la propriété que, pour $a \in A$ et $x \in I$, on a toujours $ax \in I$.

Exemple 77. Fixons $x_0 \in A$. Alors on note

$$(x_0) = \{ax_0 \mid a \in A\},$$

l'ensemble des multiples de x_0 (à gauche). C'est un idéal de A, et on dit que c'est *l'idéal principal engendré par* x_0. C'est donc un module de type fini (l'élément x_0 est une famille génératrice à lui tout seul). On le note aussi parfois Ax_0 ou, dans le cas où A est commutatif, x_0A.

Exemple 78. Prenons $A = \mathbb{Z}$. Alors chaque sous-groupe (= sous-module = idéal) de $\mathbb{Z}$ est de la forme $n\mathbb{Z}$ pour un $n \geq 1$; c'est donc un idéal principal. Vous savez peut-être que la même chose est vraie avec $A = \mathbb{K}[X]$: tout idéal de cet anneau est principal, et nous reviendrons sur ce phénomène dans le chapitre suivant.

Il y a des anneaux qui n'ont pas cette propriété. Par exemple, prenons $A = \mathbb{Z}[X]$, et

$$I = A \cdot 5 + A \cdot X ;$$

c'est une somme de deux idéaux principaux, donc en clair

$$I = \{a \cdot 5 + b \cdot X \mid a, b \in A\}.$$

Alors c'est un exercice facile que de montrer que I n'est pas principal.

Si I et J sont deux idéaux de l'anneau A, alors on peut parler de l'idéal $I+J$ (comme dans l'exemple précédent d'ailleurs), puisqu'on connait les sommes de sous-modules. Mais on peut aussi parler de IJ :

Définition 79. Le *produit* IJ des idéaux I et J est par définition l'idéal *engendré par* les éléments de la forme xy avec $x \in I$ et $y \in J$, c'est-à-dire que c'est le plus petit idéal qui contient ces éléments. Plus concrètement, IJ est l'ensemble des éléments de la forme

$$\sum_{k=1}^{m} x_k y_k$$

où m est un entier, et avec $x_k \in I$ et $y_k \in J$ pour k entre 1 et m.

Exemple 80. Si $I = (x)$ et $J = (y)$ sont des idéaux principaux, avec A commutatif, alors par définition on a $IJ = (xy)$ (vérifiez-le).

Plus généralement, si I est un idéal et M est un A-module quelconque, on peut définir IM, et on vous laisse deviner.

§2.4 Quotients

Introduction

Nous allons donner la définition du quotient V/U lorsque V est un A-module, et $U \subset V$ un sous-module. Nous allons voir qu'il y a un homomorphisme *surjectif* $V \longrightarrow V/U$ dont le noyau *est précisément* U. Ainsi, V/U est le module que l'on obtient « en remplaçant les éléments de U par des 0 ». Ça peut être très utile, quand les éléments de U ne sont pas intéressants pour le raisonnement que l'on est en train de faire, sachant que V/U peut parfois être beaucoup plus petit (et facile à comprendre) que V lui même.

Avec $V = \mathbb{Z}$ et $U = n\mathbb{Z}$, nous retrouverons bel et bien $\mathbb{Z}/n\mathbb{Z}$! Et vous êtes normalement convaincus que, par exemple, lorsqu'on travaille avec des entiers dont seule la parité nous intéresse, il est bien plus facile de se placer dans $\mathbb{Z}/2\mathbb{Z}$ qui n'a que deux éléments, plutôt que de raisonner dans $\mathbb{Z}$.

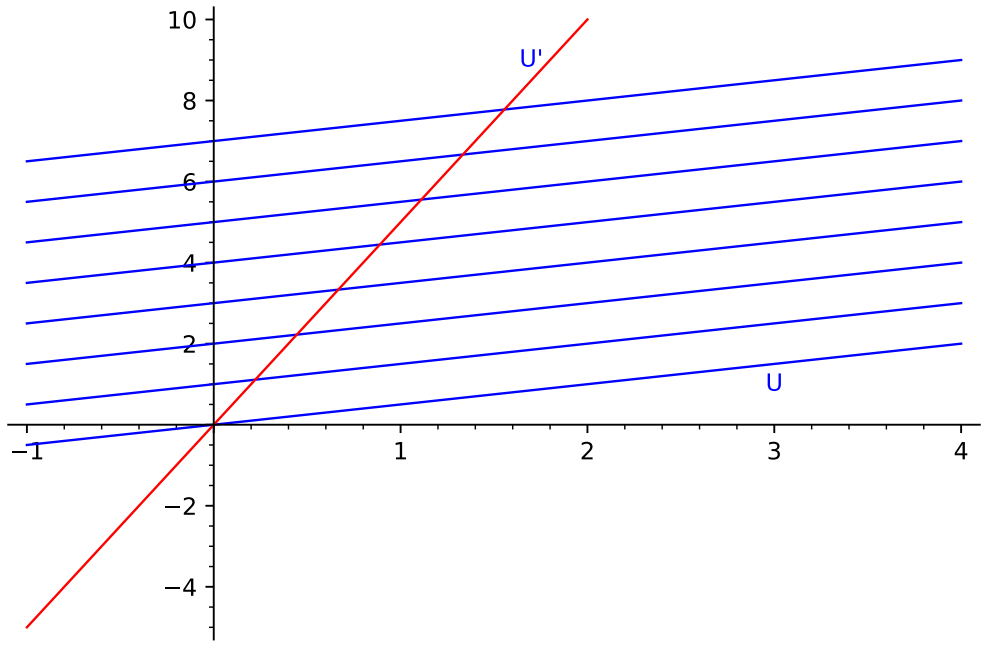
Commençons par un exemple informel, qui va motiver la définition (laquelle peut surprendre). Prenons $\mathbb{K} = \mathbb{R}$ et $V = \mathbb{R}^2$. Pour U, prenons une droite vectorielle quelconque. Comment peut-on espérer construire un module V/U avec une application linéaire $V \longrightarrow V/U$ dont le noyau serait U ? Dans ce cas précis, c'est facile. On peut prendre un U' tel que $V = U \oplus U'$ (cf chapitre précédent). On considère alors la projection

$$p : V = U \oplus U' \longrightarrow U'$$

$$p(u + u') = u' ,$$

ce qui a un sens car tout vecteur de V s'écrit de manière unique $u+u'$ avec $u \in U,\ u' \in U'$. Il est alors immédiat que le noyau de p est bien U. Donc U' peut être pris comme modèle pour V/U, avec l'homomorphisme $p : V \longrightarrow U'$. Par contre, U' n'est pas unique du tout, et c'est un peu curieux.

Or, il y a une remarque « géométrique » à faire. Faisons un dessin :



On a pris une droite U' arbitraire. Ce qui se voit bien, c'est que chaque droite (affine) parallèle à U coupe U' en un point et un seul ; en fait si $u' \in U'$, la droite

$$u' + U = \{u + u' \mid u \in U\}$$

est parallèle à U, et coupe U' en u'. *Il y a ainsi une bijection entre les éléments de U' et les droites parallèles à U.*

Pour donner une définition de V/U sans avoir à choisir une droite U', nous allons *considérer l'ensemble des droites parallèles à U, et mettre une structure d'espace vectoriel dessus.*

Passons maitenant à la version rigoureuse. Un mot d'avertissement, quitte à vous faire un peu peur : l'expérience montre que les étudiants ont des difficultés avec la notion de quotient dans le cadre le plus abstrait. Cette construction sera pourtant cruciale en M2 (s'il y a un minimum d'algèbre dans la thématique), mais la bonne nouvelle pour l'instant, c'est que dans la suite du cours du M1, nous n'aurons besoin que d'un seul exemple ou presque. Il s'agit du cas où $A = V = \mathbb{K}[X]$, et où $U = (D)$, l'idéal engendré par un polynôme D. L'objet $\mathbb{K}[X]/(D)$, qui est un $\mathbb{K}[X]$-module mais aussi un anneau, comme nous le verrons, va revenir sans cesse dans la suite.

Définition des quotients

Dans cette partie, A désigne toujours un anneau, et $\mathbb{K}$ un corps commutatif.

Définition 81. Soit V un A-module et $U \subset V$ un sous-module. Pour $v \in V$, on note

$$v + U := \{v + u \mid u \in U\} .$$

De plus, on note

$$V/U := \{v + U \mid v \in V\} .$$

(Formellement V/U est un donc un ensemble d'ensembles, chacun de la forme $v+U$, de même que ci-dessus on avait vu que V/U, sur un exemple, était identifié avec un ensemble de droites.)

Enfin, lorsque V et U sont fixés une fois pour toutes, on peut employer la notation

$$\overline{v} = v + U .$$

Avec cette notation

$$V/U = \{\overline{v} \mid v \in V\} .$$

Lemme 82. *Il existe un structure de A-module sur V/U, et une seule, telle que l'application*

$$p : V \longrightarrow V/U$$

définie par $p(v) = \overline{v}$ est un homomorphisme de A-modules.

De plus, $\ker(p) = U$.

Démonstration. Si X et Y sont des parties de V, définissons

$$X + Y = \{x + y \mid x \in X, y \in Y\} .$$

Examinons ceci dans le cas où $X = v+U$ et $Y = w+U$. On constate rapidement que

$$X + Y = (v + U) + (w + U) = (v + w) + U .$$

Ceci nous donne une opération + sur V/U, et il est immédiat que $\overline{v}+\overline{w} = \overline{v + w}$. De plus, puisque p est surjective, il est très simple de vérifier que + donne bien

une structure de groupe abélien sur V/U. Par exemple, admettons que l'on veuille vérifier la commutativité, c'est-à-dire que $x+y = y+x$ pour $x, y \in V/U$. Choisissons $v \in V$ tel que $\overline{v} = x$, et de même prenons w tel que $\overline{w} = y$, alors

$$x + y = \overline{v} + \overline{w} = \overline{v + w} = \overline{w + v} = \overline{w} + \overline{v} = y + x\,.$$

Nous avons utilisé la commutativité de la loi + sur V. Notons que l'élément neutre est $\overline{0} = U$.

On fait pareil avec les autres propriétés à vérifier (associativité...). Et sur le même modèle, on définit, pour $a \in A$ et X une partie de V, la partie aX par

$$aX = \{ax \mid x \in X\}\,.$$

Si $X = v + U$ alors $aX = av + U$. Ceci nous donne une opération $A \times V/U \longrightarrow V/U$ qui complète la structure de A-module (les vérifications étant là encore très simples). On a $\overline{av} = a\overline{v}$, par définition.

L'unicité provient du fait que p est surjective (on vous laisse vérifier ceci).

Nous devons maintenant examiner le noyau de p. Il s'agit des $v \in V$ tels que $p(v) = \overline{0} = \overline{v}$, ce qui par définition signifie $U = v + U$. Clairement, ceci arrive si et seulement si $v \in U$. □

Proposition 83. *Soit U, V comme ci-dessus et $p : V \longrightarrow V/U$ l'application quotient. Soit W un autre A-module et $f : V \longrightarrow W$ un homomorphisme. On suppose que $f(u) = 0$ pour tous les $u \in U$.*

Alors il existe un unique homomorphisme $\overline{f} : V/U \longrightarrow W$ tel que $f = \overline{f} \circ p$. Ou ce qui revient au même : pour $v \in V$ on a $f(v) = \overline{f}(\overline{v})$.

Démonstration. Soit $x \in V/U$. On peut choisir un $v \in V$ tel que $x = p(v) = \overline{v}$. Ce v n'est pas unique, mais par contre, l'élément $f(v)$ ne dépend pas du choix : en effet, si $p(v') = p(v) = x$, alors $u = v - v' \in \ker(p) = U$, donc $f(u) = 0 = f(v) - f(v')$. On peut donc poser $\overline{f}(x) = f(v)$ pour un v quelconque tel que $x = p(v)$, et ceci est bien défini. On a $f(v) = \overline{f}(p(v))$ par définition.

Il faut vérifier que $\overline{f}$ est un homomorphisme, mais c'est très facile, par exemple si $x = \overline{v}$ et $y = \overline{w}$ alors $x + y = \overline{v + w}$ de sorte que

$$\overline{f}(x + y) = \overline{f}(\overline{v + w}) = f(v + w) = f(v) + f(w) = \overline{f}(\overline{v}) + \overline{f}(\overline{w}) = \overline{f}(x) + \overline{f}(y)\,.$$

Et ainsi de suite. □

Corollaire 84. *Soit $f : V \longrightarrow W$ un homomorphisme surjectif entre A-modules, et soit $U = \ker f$. Alors l'application induite $\overline{f} : V/U \longrightarrow W$ est un isomorphisme.*

On résume parfois ce corollaire en écrivant

$$V/\ker f \cong \mathrm{Im}(f)\,.$$

Démonstration. Puisque f est surjective, tout $w \in W$ est de la forme $w = f(v) = \overline{f}(\overline{v})$, donc $\overline{f}$ est surjective. De plus, si $x \in \ker \overline{f}$, alors en prenant $v \in V$ tel que $x = \overline{v}$ on a $\overline{f}(x) = f(v) = 0$, d'où $v \in U$ et $\overline{v} = \overline{0} = x$. On a bien $\ker \overline{f} = \{\overline{0}\}$, donc $\overline{f}$ est également injective. □

Exemple 85. Nous avons promis que pour $V = \mathbb{Z}$ et $U = n\mathbb{Z}$ on retrouvait $\mathbb{Z}/n\mathbb{Z}$ comme on le connaît. Montrons-le : soit M le groupe abélien que vous avez appelé $\mathbb{Z}/n\mathbb{Z}$ les années précédentes ; tout le monde n'a peut-être pas eu la même définition, mais vous nous accorderez qu'il y a un homomorphisme surjectif $f : \mathbb{Z} \longrightarrow M$ dont le noyau est $n\mathbb{Z}$. Par le corollaire, $M \cong \mathbb{Z}/n\mathbb{Z}$, où l'écriture $\mathbb{Z}/n\mathbb{Z}$ désigne bien la nouvelle notion introduite dans ce chapitre.

Avant de donner d'autres exemples :

Définition 86. Soit U un sous-module de V, et soit $E \subset V$ un sous-ensemble – on ne suppose pas que E est un sous-module, en général. On dit que U et E sont *supplémentaires* lorsque tout $v \in V$ peut s'écrire de manière unique $v = u + e$ avec $u \in U$ et $e \in E$.

Bien sûr, si $V = U \oplus U'$, alors $E = U'$ est un supplémentaire de U, mais il y a d'autres exemples. Par exemple, si $A = V = \mathbb{Z}$, $U = n\mathbb{Z}$, et $E = \{0, 1, 2, \ldots, n-1\}$.

Lemme 87. *Soit U un sous-module de V, et soit E un supplémentaire de U. Alors l'application quotient*

$$p : V \longrightarrow V/U$$

donne une bijection entre E et V/U. Si E est un sous-module, alors p est un isomorphisme de modules ; si A est une algèbre sur $\mathbb{K}$ et si E est un $\mathbb{K}$-espace vectoriel, alors p est un isomorphisme d'espaces vectoriels.

Démonstration. La définition même de « supplémentaire » rend évident le fait que tout $x \in V/U$ s'écrit $x = p(e)$ pour un $e \in E$ unique. Le reste provient du fait que p est un homomorphisme de modules, donc sa restriction à E (si E est aussi un module) est encore un homomorphisme. □

Exemple 88. On retrouve bien sûr le fait que

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$$

(et le fait que les n éléments à droite sont distincts). Voici un autre exemple : prenons $A = \mathbb{K}[X] = V$, et $U =$ les multiples de X^3 comme dans l'introduction. On pose alors $E = \mathcal{V}ect(1, X, X^2)$. C'est un sous-espace vectoriel de V, mais pas un sous-module (si on multiplie un élément de E par X, on ne tombe pas forcément dans E !). Le lemme s'applique, car U et E sont supplémentaires. Ceci montre que V/U est isomorphe comme espace vectoriel à E, donc il est bien de dimension 3 avec pour base $\overline{1}, \overline{X}, \overline{X^2}$.

Quotient d'un anneau par un idéal

Nous venons de voir les quotients de modules. Voyons les quotients dans le monde des anneaux (dans un autre chapitre, nous verrons les quotients de groupes non-commutatifs).

Lemme 89. *Soit I un idéal de l'anneau A. On suppose que A est commutatif. Alors le A-module A/I est également un anneau commutatif, et l'application naturelle $p : A \longrightarrow A/I$ est un homomorphisme d'anneaux.*

Enfin, si A est une $\mathbb{K}$-algèbre, où $\mathbb{K}$ est un corps commutatif, alors A/I aussi, et l'application p est un homomorphisme d'algèbres.

Démonstration. Soient $x, y \in A/I$. On choisit $a, b \in A$ tels que $p(a) = x$ et $p(b) = y$; montrons que l'élément $p(ab) = ab + I \in A/I$ ne dépend pas du choix de a ou b, mais seulement de x et y.

En effet, si $p(a') = p(a) = x$ et $p(b') = p(b) = y$, alors $i = a' - a \in \ker(p) = I$ et de même $j = b' - b \in I$. Par suite

$$a'b' + I = ab + (ib' + aj + ij) + I = ab + I\,,$$

la dernière égalité provenant du fait que $ib' + aj + ij \in I$, car I est un idéal, et A est commutatif (c'est crucial !), donc $ib' = b'i \in I$.

On peut donc définir $xy = ab + I$, et ceci a un sens. Ceci donne une multiplication sur A/I, et $p(a)p(b) = p(ab)$ par définition. On en déduit facilement que A/I est un anneau commutatif.

La démonstration du « enfin » vous est laissée en exercice. □

Exemple 90. On retrouve le fait que $\mathbb{Z}/n\mathbb{Z}$ est un anneau, ainsi que la formule $\overline{ab} = \overline{a}\overline{b}$.

Lemme 91. *Soit $f : A \longrightarrow B$ un homomorphisme d'anneaux, et I un idéal de A tel que $f(I) = \{0\}$. Alors l'application induite $\overline{f} : A/I \longrightarrow B$ est un homomorphisme d'anneaux.*

De plus, si f est surjective et $I = \ker f$, alors l'isomorphisme $A/I \cong B$ induit par f est un isomorphisme d'anneaux.

Démonstration. Exercice. □

Dans la suite, nous allons nous concentrer sur un type d'exemple très important : on prend un corps commutatif $\mathbb{K}$, l'anneau $A = \mathbb{K}[X]$, et l'idéal $I = (D)$ principal engendré par le polynôme non-nul $D = a_0 + a_1X + \cdots + a_dX^d$. En tant que A-module, le cas de A/I est crucial, comme nous le verrons dans le chapitre suivant (en gros : tout $\mathbb{K}[X]$-module finiment engendré est un produit de modules de la forme $\mathbb{K}[X]/(D)$). En tant qu'anneau, le cas de A/I sera étudié à la loupe au second semestre.

Posons donc

$$E = \{P \in \mathbb{K}[X] \mid \deg(P) < d\}.$$

Comme vous le savez, dans $\mathbb{K}[X]$ on peut faire des divisions euclidiennes, c'est-à-dire que pour tout $P \in \mathbb{K}[X]$, on peut écrire

$$P = DQ + R$$

avec $\deg(R) < d$, et ceci de manière unique. Ceci revient à dire que (D) et E sont supplémentaires dans $\mathbb{K}[X]$ (pensez-y !). D'après un lemme ci-dessus, on en déduit que A/I est isomorphe, comme $\mathbb{K}$-espace vectoriel, à E. Plus

précisément, en écrivant comme d'habitude $P \mapsto \overline{P}$ pour l'application $A \longrightarrow A/I$, les éléments $\overline{1}, \overline{X}, \ldots, \overline{X^{d-1}}$ forment une base de A/I.

Nous allons poser $x = \overline{X}$. On a $\overline{X^k} = x^k$ puisque l'application $A \longrightarrow A/I$ est un homomorphisme d'anneaux. Finalement, tout élément de A/I peut s'écrire de manière unique

$$\lambda_0 + \lambda_1 x + \lambda_2 x^2 + \cdots + \lambda_{d-1} x^{d-1} .$$

Et enfin, pour faire des calculs dans A/I, il suffit de se rappeler que $D(x) = 0$. En effet $D(x) = \overline{D(X)}$, et bien sûr $D \in (D)$ donc $\overline{D} = 0$. Nous avons tout ce qu'il nous faut pour travailler dans cet anneau quotient.

Exemple 92. Prenons $\mathbb{K} = \mathbb{R}$ et $D = X^2 + 1$. Appelons $K = \mathbb{R}[X]/(X^2 + 1)$. Comme espace vectoriel sur $\mathbb{R}$, cet anneau est de dimension 2, avec comme base $1, x$. De plus, on a la relation $x^2 + 1 = 0$. En écrivant ça $x^2 = -1$, on peut s'amuser à multiplier $z = a + bx$ par $w = a' + b'x$:

$$zw = aa' - bb' + (ab' + a'b)x .$$

Évidemment on se dit qu'on a probablement $K \cong \mathbb{C}$, et il est très facile de le montrer. Soit $\varphi\colon \mathbb{R}[X] \longrightarrow \mathbb{C}$ l'homomorphisme $P \mapsto P(i)$. Alors φ est nul sur l'idéal $I = (X^2 + 1)$, puisque

$$\varphi((X^2 + 1) \cdot P) = (i^2 + 1)\varphi(P) = 0 .$$

On a donc un homomorphisme de $\mathbb{R}$-algèbres induit $\overline{\varphi}\colon \mathbb{R}[X]/(X^2+1) \longrightarrow \mathbb{C}$, qui est évidemment surjectif, et puisqu'il est $\mathbb{R}$-linéaire entre deux espaces de même dimension, c'est un isomorphisme (qui envoie x sur i) On a bien $K \cong \mathbb{C}$.

Au passage, on a même un peu plus. Soit $\varphi_-\colon \mathbb{R}[X] \longrightarrow \mathbb{C}$ défini par $P \mapsto P(-i)$. Comme ci-dessus, on démontre que $\overline{\varphi}_-$ est un isomorphisme entre K et $\mathbb{C}$. On voit alors que $\overline{\varphi}_- \circ \overline{\varphi}^{-1}\colon \mathbb{C} \longrightarrow \mathbb{C}$ est un automorphisme (de $\mathbb{R}$-algèbres) qui envoie i sur $-i$. Il s'agit bien sûr de la conjugaison complexe.

Exemple 93. Essayons avec $\mathbb{K} = \mathbb{C}$ et $D = X^2 + 1$. On a maintenant $X^2 + 1 = (X + i)(X - i)$, et

$$\frac{1}{2i}(X + i) - \frac{1}{2i}(X - i) = 1 .$$

Les idéaux $I = (X + i)$ et $J = (X - i)$ vérifient donc $I + J = \mathbb{K}[X]$. D'après le lemme chinois (cf les exercices), on a

$$\mathbb{C}[X]/(X^2 + 1) \cong \mathbb{C}[X]/(X + i) \times \mathbb{C}[X]/(X - i)\,.$$

Par ailleurs, l'anneau $\mathbb{C}[X]/(P)$, avec P quelconque de degré 1, est une algèbre sur $\mathbb{C}$ de dimension 1 : c'est donc simplement $\mathbb{C}$! Finalement $\mathbb{C}[X]/(X^2+1) \cong \mathbb{C} \times \mathbb{C}$. Ce n'est pas un corps.

Exemple 94. Et maintenant, prenons $\mathbb{K} = \mathbb{Q}$ et $D = X^2 - 2$. L'anneau $K = \mathbb{Q}[X]/(X^2 - 2)$ est une algèbre de dimension 2 sur $\mathbb{Q}$, avec pour base $1, x$, et $x^2 - 2 = 0$, donc $x^2 = 2$. Les calculs dans K sont du type :

$$(a + bx)(a' + b'x) = aa' + 2bb' + (ab' + a'b)x\,.$$

Montrons que $K \cong \mathbb{Q}[\sqrt{2}]$ (cf les exercices de la feuille 1 pour la notation). On note $\varphi_\pm : \mathbb{Q}[X] \longrightarrow \mathbb{Q}[\sqrt{2}]$ l'homomorphisme $P \mapsto P(\pm\sqrt{2})$. Alors $\varphi_\pm$ vaut 0 sur l'idéal engendré par $X^2 - 2$, et on a donc un homomorphisme induit $\overline{\varphi}_\pm : \mathbb{Q}[X]/(X^2 - 2) \longrightarrow \mathbb{Q}[\sqrt{2}]$; ce dernier est surjectif, et en comparant les dimensions, on constate que $\overline{\varphi}_\pm$ est un isomorphisme. Il envoie x sur $\pm\sqrt{2}$.

On note que $\overline{\varphi}_- \circ \overline{\varphi}_+^{-1}$ est un automorphisme de $\mathbb{Q}[\sqrt{2}]$ qui envoie $\sqrt{2}$ sur $-\sqrt{2}$. Il n'est pas du tout trivial qu'un tel automorphisme existe !

Chapitre 3

Anneaux euclidiens et leurs modules

Dans ce chapitre, nous allons définir de nombreux types d'anneaux : les anneaux intègres, principaux, factoriels (dans les exos), et noethériens. Mais ce qui nous intéresse vraiment, ce sont les anneaux euclidiens, qui vérifient tout ça à la fois (nous allons donc voir qu'un anneau euclidien est intègre, principal, factoriel et noethérien !). Le vocabulaire est surtout là pour que vous l'ayez vu une fois, mais ce n'est pas un objectif en soi.

Après avoir introduit les anneaux euclidiens, nous allons donner un théorème (très fort !) de classification de leurs modules. Dans le cas de $\mathbb{K}[X]$, on retrouve les théorèmes classiques de réduction des endomorphismes – et bien plus.

§3.1 Anneaux euclidiens

Définitions

Définition 95. Un anneau A est dit *intègre* s'il est commutatif, et si pour $a, b \in A$ on a $ab = 0 \implies a = 0$ ou $b = 0$.

Exemple 96. La plupart des anneaux commutatifs que vous connaissez sont intègres, donc il est plus instructif de voir un contre-exemple : l'anneau $\mathbb{Z}/4\mathbb{Z}$

n'est pas intègre, puisque $\overline{2} \neq 0$ mais $\overline{2} \cdot \overline{2} = \overline{0}$.

Remarque. Très souvent, on utilise cette propriété sous la forme : si $ax = ay$, dans un anneau intègre, avec $a \neq 0$, alors $x = y$ (« on a le droit de simplifier »). En effet, on a $a(x - y) = 0$ donc $x - y = 0$.

Définition 97. Un anneau A est dit *euclidien* s'il est intègre, et s'il existe une fonction

$$v : A - \{0\} \longrightarrow \mathbb{N}$$

telle que :

1. $v(a) \leq v(ab)$ pour tous $a, b \in A$ non nuls,
2. pour tout $a, b \in A$, avec $b \neq 0$, il existe $q, r \in A$ tels que

$$a = bq + r$$

et avec $v(r) < v(b)$, ou alors $r = 0$.

On pose parfois $v(0) = -\infty$, pour que v soit définie partout. Notons que l'on n'exige pas, dans cette définition, que q et r soient uniques.

Exemple 98. Il y a deux exemples fondamentaux. Tout d'abord $A = \mathbb{Z}$, avec $v(x) = |x|$. Les propriétés ci-dessus vous sont familières dans ce cas, et en plus les éléments q et r sont uniques si on exige $r \geq 0$. On les appelle le *quotient* et le *reste* dans la *division euclidienne* de a par b.

Autre exemple fondamental, $A = \mathbb{K}[X]$ avec $v(P) = \deg(P)$. Là encore, vous avez déjà vu tout ça, ainsi que le fait que le quotient et le reste sont uniques.

Il y a un exemple dégénéré : supposons que $A = \mathbb{K}$ soit un corps, et posons $v(x) = 0$ pour $x \neq 0$. Alors la propriété (1) est satisfaite, et pour la (2), il suffit de prendre $r = 0$ et $q = \frac{1}{b}$. Donc les corps (commutatifs) sont des cas particuliers d'anneaux euclidiens, et ce que nous allons montrer dans ce chapitre et surtout dans le suivant va s'appliquer aux corps. Mais ça n'a pas beaucoup d'intérêt : les résultats obtenus sont tous déjà connus dans le cas des corps.

Définition 99. Un anneau A est dit *principal* lorsqu'il est intègre, et que ses idéaux sont tous principaux. On rappelle qu'un idéal I est dit principal lorsqu'il est de la forme $I = (x) = Ax$ (= les multiples de x).

Lemme 100. *Un anneau euclidien est principal.*

Démonstration. Soit I un idéal de l'anneau euclidien A. Si $I = \{0\}$ alors $I = (0)$ est principal, donc un peut supposer que $I \neq \{0\}$ et poser

$$m = \min\{v(x) \mid x \in I, x \neq 0\}$$

puis prendre un $b \in I$ non-nul tel que $v(b) = m$. On a certainement $(b) \subset I$.

Pour $a \in I$ quelconque, on écrit une division euclidienne

$$a = bq + r\,.$$

On a $r = a - bq \in I$. Si on avait $r \neq 0$, alors on aurait $v(r) < v(b) = m$ ce qui est absurde. Donc $r = 0$ et $a = bq \in (b)$, d'où $I = (b)$. □

Si $I = (x)$, l'élément x est-il unique ? Une définition d'abord :

Définition 101. Si A est un anneau quelconque, on note

$$A^\times = \{a \in A \mid a^{-1} \text{ existe } \}\,.$$

Alors $A^\times$ est un groupe pour la multiplication, et ses éléments s'appellent les *unités* de A (on peut dire évidemment aussi les *éléments inversibles* de A).

Exemple 102. Par exemple $\mathbb{Z}^\times = \{\pm 1\}$ et $\mathbb{K}[X]^\times = \mathbb{K}^\times = \mathbb{K} \setminus \{0\}$.

Lemme 103. *Soit A un anneau intègre (par exemple un anneau euclidien), et soient $x, y \in A$ tels que $(x) = (y)$. Alors $y = xu$ pour un $u \in A^\times$.*

Démonstration. Puisque $y \in (y) = (x)$, on a $y = xu$ pour un $u \in A$. Mais on a aussi $x = yv$ pour un $v \in A$, de la même manière, de sorte que $y = xu = yvu$. Dans un anneau intègre, on peut simplifier, et obtenir $vu = 1$, donc $u \in A^\times$. □

Les éléments *irréductibles* et *premiers* vont jouer un rôle important dans la suite :

Définition 104. Soit A un anneau commutatif et $p \in A$ non inversible. On dit que p est *irréductible* lorsque, pour $a, b \in A$, l'égalité $p = ab$ entraîne que a ou b est une unité.

Par exemple, dans $\mathbb{K}[X]$, on retrouve la définition de polynôme irréductible. Notez que, dans ce cas comme dans le cas général, la décision d'exclure les éléments inversibles n'est qu'une convention (un polynôme constant n'est pas appelé « irréductible », c'est comme ça), qui s'avère bien pratique à l'usage.

Définition 105. Soit A un anneau commutatif et $p \in A$ non inversible. On dit que p est *premier* lorsque, à chaque fois que p divise ab avec a et b dans A, alors p divise a ou p divise b.

Ces deux définitions sont en fait très proches :

Lemme 106. *Soit A un anneau commutatif et $p \in A$ non inversible. Si A est intègre, alors si p est premier, il est également irréductible. Si A est principal, la réciproque est vraie, c'est le lemme de Gauss : si p est irréductible, il est aussi premier.*

Démonstration. Sous forme d'exercices (pour la deuxième partie, il vaut mieux lire ci-dessous les rappels sur les pgcd, ça donne du vocabulaire pratique, mais insistons sur le fait qu'on peut faire la démonstration tout de suite). □

Dans notre cours, tous les anneaux considérés seront en général principaux (et même euclidiens), et donc on ne va pas s'embarrasser trop de la différence entre « irréductible » et « premier ».

Exemple 107. Dans $\mathbb{Z}$, on réalise bien vite qu'un irréductible est un élément p dont les seuls diviseurs sont $1, -1, p$ et $-p$. On retrouve la définition d'un nombre premier (sauf que vous n'étiez peut-être pas habitués à dire que -2 est un nombre premier).

Définition 108. Soit A un anneau commutatif et $I \subset A$ un idéal. On dit que I est un *idéal premier* lorsque, pour $a, b \in A$, l'égalité $ab \in I$ entraîne que $a \in I$ ou $b \in I$.

Voici un mini-exo que vous devez faire tout de suite : I est un idéal premier si et seulement si l'anneau quotient A/I est intègre. (Ce n'est qu'une affaire de définitions.)

Dans le cas qui nous intéresse, celui des anneaux euclidiens, les choses se simplifient :

Proposition 109. *Soit A un anneau principal (par exemple un anneau euclidien), et $I = (p)$ un idéal quelconque. Alors les conditions suivantes sont équivalentes :*

1. *A/I est un corps ;*
2. *A/I est intègre ;*
3. *I est un idéal premier ;*
4. *p est un élément premier.*

Démonstration. Un corps est intègre, donc (1) $\implies$ (2) est toujours vrai, et vous venez de montrer que (2) $\iff$ (3). Voyons pourquoi (3) $\implies$ (4) : soient $a, b \in A$ tels que $p = ab$, de sorte que $ab \in I$. Si $a \in I$, on peut écrire $a = px$ avec $x \in A$, et $p = ab = pxb$, d'où $xb = 1$, et on voit que b est inversible. De même si $b \in I$, on voit que a est inversible. On a bien montré que p est irréductible (donc premier).

Jusqu'ici, on n'a pas utilisé le fait que A est principal. Mais ce qui nous intéresse particulièrement, c'est donc l'implication (4) $\implies$ (1), pour laquelle l'hypothèse va nous servir.

Montrons donc que A/I est un corps, en supposant que p est premier. D'après un exo de la feuille précédente, ceci revient à montrer que les seuls idéaux de A/I sont $\{\overline{0}\}$ et A/I ; et d'après un autre exo, ceci revient à montrer que, si J est un idéal de A tel que $I \subset J \subset A$, alors on a $J = I$ ou $J = A$.

Montrons ceci, donc soit J tel que $I \subset J \subset A$. Puisque A est principal, il existe un élément q tel que $J = (q)$. De l'inclusion $(p) \subset (q)$, on tire $p = qa$ avec $a \in A$. Puisque p est premier, il y a deux possibilités. Soit a est une unité, et alors de $q = pa^{-1} \in (p)$ on tire $(q) \subset (p)$, et donc $J = I$. Soit q est une unité, et alors $J = A$: en effet, tout $x \in A$ vérifie alors $x = xq^{-1}q \in (q)$. □

Dans un anneau euclidien (ou même principal), on peut développer la théorie des pgcd comme on le fait dans $\mathbb{Z}$ ou dans $\mathbb{K}[X]$. Comme vous l'avez déjà fait dans ces deux cas-là, et comme nous ne verrons pas d'autres exemples intéressants d'anneaux euclidiens, on va se permettre de faire ça un peu vite.

Définition 110. Soit A un anneau euclidien, et soient $a_1, \ldots, a_n \in A$. On note

$$(a_1, \ldots, a_n) = Aa_1 + \cdots + Aa_n$$

$$= \{u_1 a_1 + \cdots + u_n a_n \mid u_i \in A\},$$

qui est le plus petit idéal de A contenant les a_i. Tout élément $d \in A$ vérifiant

$$(a_1, \ldots, a_n) = (d)$$

est appelé un *pgcd* de $a_1, \ldots, a_n$. Un pgcd existe toujours, et est défini à un inversible près (lemme 103). Par définition, puisque $d \in (d)$, il existe des éléments u_i tels que

$$u_1 a_1 + \cdots + u_n a_n = d,$$

ce qu'on appelle une *relation de Bézout.*

Enfin, si 1 est un pgcd des éléments $a_1, \ldots, a_n$, on dit qu'ils sont *premiers entre eux.*

§3.2 Anneaux noethériens

Soit donc V un module de type fini sur l'anneau A. Par définition, il existe un entier n et une application surjective

$$\pi : A^n \longrightarrow V.$$

Si $K = \ker(\pi)$, nous avons donc

$$V \cong A^n / K.$$

Mais que dire de K ? La première question est certainement : est-ce K est de type fini ? En fait, la réponse dépend de l'anneau A.

Définition 111. On dit que l'anneau A est *noethérien* si tout sous-module d'un A-module de type fini est encore de type fini.

Il existe des variantes de cette définition, et pour cause :

Proposition 112. *Soit A un anneau commutatif. Alors les conditions suivantes sont équivalentes :*

1. *A est noethérien ;*

2. *tout idéal de A est de type fini comme module, c'est-à-dire que tout idéal est de la forme $(a_1, \dots, a_k)$ avec des $a_i \in A$.*
3. *toute suite croissante d'idéaux de A*

$$I_0 \subset I_1 \subset I_2 \subset \cdots$$

est stationnaire (il existe k tel que $I_\ell = I_k$ pour tous les $\ell \geq k$).

Démonstration. Puisqu'un idéal I est un module, l'implication (1) $\implies$ (2) est triviale ; et (2) $\implies$ (3) est presque triviale aussi, puisque si on pose $I = \bigcup_k I_k$, alors I est lui-même un idéal, donc il est de la forme $I = (a_1, \dots, a_n)$, et il suffit de prendre k suffisamment grand pour que $a_1, \dots, a_n \in I_k$ pour obtenir $I_k = I$.

Il n'y a pas beaucoup de travail non plus à faire pour obtenir (3) $\implies$ (2). En effet prenons un idéal I, et posons $I_0 = (0)$. Si I est non-nul, soit $a_1 \in I$ avec $a_1 \neq 0$ et $I_1 = (a_1)$. Si $I \neq I_1$, on prend $a_2 \in I \setminus I_1$ et on pose $I_2 = (a_1, a_2)$. Si $I \neq I_2$, on prend $a_3 \in I \setminus I_2$ et on pose $I_3 = (a_1, a_2, a_3)$. Il serait absurde que l'on puisse continuer ainsi à l'infini, puisque l'on est en train de construire une suite $I_0 \subset I_1 \subset I_2 \subset \cdots$ d'idéaux avec $I_k \neq I_{k+1}$, et on a fait l'hypothèse (3) qui affirme que ça n'arrive jamais. Donc à une certaine étape, on a $I = I_k = (a_1, \dots, a_k)$.

La partie intéressante est (2) $\implies$ (1) – le fait qu'il suffit de regarder les idéaux pour savoir que A est noethérien. Soit donc A un anneau commutatif qui satisfait (2), et commençons par prendre un sous-module $U \subset A^n$ pour un certain n. On va montrer que U est de type fini. On procède par récurrence sur n, le cas $n = 1$ étant l'hypothèse (2).

On suppose donc le résultat pour $n - 1$. Soit $p : A^n \longrightarrow A$ la projection

$$p(x_1, \dots, x_n) = x_n \,.$$

L'image $p(U) \subset A$ est un idéal. Par hypothèse, on a donc $p(U) = (a_1, \dots, a_k)$ pour $a_i \in A$, et par construction on peut choisir $u_i \in U$ tel que $p(u_i) = a_i$. Si U' est le sous-module de U engendré par les u_i, alors on a

$$U = U' + \ker(p) \cap U$$

(par contre la somme n'est pas directe en général). En effet, si $u \in U$ vérifie $p(u) = \sum t_i a_i$ avec $t_i \in A$, alors $p(u - \sum t_i u_i) = 0$, donc $u - \sum t_i u_i \in \ker(p) \cap U$.

Mais $\ker(p) \cap U$ est un sous-module de $A^{n-1} \times \{0\}$, que l'on identifie à A^{n-1}. Par récurrence, $\ker(p) \cap U$ est engendré par, disons, $w_1, \ldots, w_r$, donc finalement U est engendré par $w_1, \ldots, w_r, u_1, \ldots, u_k$. On a bien montré que U est de type fini.

Le cas général va suivre tout de suite de ce cas particulier : soit V un module de type fini, soit $\pi : A^n \longrightarrow V$ surjective (qui existe par définition), soit $W \subset V$, et soit $U = \pi^{-1}(W)$. Alors U est de type fini (on vient de le démontrer), donc $W = p(U)$ aussi. □

Corollaire 113. *Tout anneau principal, et en particulier tout anneau euclidien, est noethérien.*

Plus précisément, si A est principal et si un module V est engendré par n éléments, alors tout sous-module $U \subset V$ peut être engendré par m éléments avec $m \leq n$.

Démonstration. On a la propriété (2), puisque tout idéal est engendré par un seul élément dans ce cas. D'après la proposition, l'anneau A est noethérien.

Mais on a mieux. En effet, regardons la démonstration de (2) $\Longrightarrow$ (1) ci-dessus : dans le cas d'un anneau principal, on peut prendre $k = 1$, le module U' est engendré par un seul élément, et par récurrence on peut supposer que $\ker(p) \cap U$ est engendré par moins de $n-1$ éléments, donc U est engendré par moins de n éléments. □

Revenons à nos observations du début de cette partie, et au A-module V de type fini, qui est donc de la forme

$$V \cong A^n/K$$

où $K = \ker(\pi)$ avec $\pi : A^n \longrightarrow V$ surjective. Si on suppose que A est euclidien, on sait que K est de type fini, donc qu'il existe

$$\varphi : A^m \longrightarrow K$$

qui est surjective, avec $m \leq n$. Dans la suite, on va utiliser la même lettre φ pour désigner l'application

$$\varphi : A^m \longrightarrow A^n$$

obtenue en composant avec l'inclusion de K dans A^n. On a donc $\varphi(A^m) = K$, et

$$V \cong A^n/\mathrm{Im}(\varphi)\,.$$

C'est donc une application φ entre deux modules libres qui contient l'information que nous cherchons : comprendre V revient à comprendre φ.

Attention, trouver explicitement une famille génératrice pour le sous-module K de A^n peut être difficile, même pour $A = \mathbb{Z}$. Cependant, le simple fait de savoir que K est de type fini va nous permettre d'obtenir un théorème très fort, aux conséquences bien concrètes.

L'exemple suivant relève, lui, un peu de l'astuce, et on pourrait s'en passer, mais le résultat est tellement simple qu'il mérite d'être mentionné.

Exemple 114. Prenons $A = \mathbb{K}[X]$ et (V, f) un A-module. On suppose que V est de dimension finie sur $\mathbb{K}$, et plus précisément, on va supposer que $V = \mathbb{K}^n$ avec la base canonique $e_1, \ldots, e_n$ dans laquelle l'endomorphisme $f\colon V \longrightarrow V$ a pour matrice F. Il y a une façon assez astucieuse de trouver φ comme ci-dessus, très explicite de surcroît, et en un sens c'est plus facile qu'avec $A = \mathbb{Z}$.

En effet, prenons d'abord pour $\pi\colon A^n \longrightarrow V$ l'application évidente, c'est-à-dire $\pi(a_1, \ldots, a_n) = a_1e_1 + \cdots a_ne_n$. Définissons alors $\varphi\colon A^n \longrightarrow A^n$ comme étant l'application A-linéaire définie par la matrice $F - XI \in M_n(A)$ (où I est la matrice identité). En d'autres termes $\varphi(v) = Fv - Xv$ en voyant $v \in A^n$ comme un vecteur-colonne. (Notons que le déterminant de l'application φ est le polynôme caractéristique de F !) Montrons que l'image de φ est le noyau de φ, comme souhaité.

En effet, on a $\pi(\varphi(v)) = \pi(Fv) - \pi(Xv)$. On va se servir d'abord uniquement du fait que π est A-linéaire, pour remarquer que $\pi(Xv) = X \cdot \pi(v) = f(\pi(v)) = F\pi(v)$, par définition de la structure de module sur V. Mais nous avons aussi $\pi(Fv) = F\pi(v)$, comme nous allons le vérifier. La remarque essentielle est celle-ci : si on voit $\mathbb{K}$ comme un sous-anneau de A, et donc $\mathbb{K}^n$ comme un sous-espace vectoriel de A^n, alors $\pi(x) = x$ pour $x \in \mathbb{K}^n$. Ainsi, si la base canonique de A^n est $\varepsilon_1, \ldots, \varepsilon_n$, nous avons $F\varepsilon_i = Fe_i =$ la i-ème colonne de F, dont les coefficients sont dans $\mathbb{K}$, et donc $\pi(F\varepsilon_i) = F\varepsilon_i = Fe_i = F\pi(\varepsilon_i)$. Comme c'est vrai pour tout i, les deux applications A-linéaires $\pi \circ F$ et $F \circ \pi\colon A^n \longrightarrow V$

doivent coïncider.

Finalement, $\pi(\varphi(v)) = 0$, donc l'image de φ est contenue dans le noyau de π. L'application $A^n \longrightarrow A^n/\ker(\pi)$ se factorise donc par $\mathrm{Im}(\varphi)$, ce qui donne une application

$$A^n/\mathrm{Im}(\varphi) \longrightarrow A^n/\ker(\pi)$$

qui est évidemment surjective. Il suffit de montrer que les deux espaces ont la même dimension comme espace vectoriels sur $\mathbb{K}$ pour conclure que cette application est un isomorphisme, et donc que $\mathrm{Im}(\varphi) = \ker(\varphi)$. Or la dimension de $A^n/\ker(\varphi) \cong V$ est n, et nous allons montrer que la dimension de $A^n/\mathrm{Im}(\varphi)$ est $\leq n$, ce qui suffit. Et en effet il suffit d'écrire

$$Xv = Fv + (Xv - Fv) \equiv Fv \bmod \mathrm{Im}(\varphi)$$

pour constater que la famille $\overline{\varepsilon}_1, \ldots, \overline{\varepsilon}_n$ engendre $A^n/\mathrm{Im}(\varphi)$ comme $\mathbb{K}$-espace vectoriel.

§3.3 La forme normale de Smith

Puisque les applications entre modules libres nous intéressent désormais, il va être utile de montrer des choses sur les matrices à coefficients dans l'anneau euclidien A.

Nous allons faire des opérations sur les lignes et les colonnes qui sont les mêmes que celles auxquelles vous êtes habitués dans le cas de l'algèbre linéaire sur un corps :

- échanger deux lignes ou deux colonnes ;
- multiplier une ligne ou une colonne par un élément *inversible* de A ;
- ajouter à une ligne un multiple d'une autre ligne (et pareil avec les colonnes).

De même que dans le cas des corps, on montre que faire une opération sur les lignes d'une matrice M revient à la multiplier à gauche par une matrice *inversible* P : par exemple pour ajouter à la ligne 1 un multiple de la ligne 2,

on prend

$$P = \begin{pmatrix} 1 & a & & & \\ 0 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

On vérifie que PM est égale à M à qui on a fait subir $L_1 \leftarrow L_1 + aL_2$. De même, les opérations sur les colonnes reviennent à multiplier par une matrice inversible, à coefficients dans A, à droite.

Théorème 115. *Soit M une matrice $n \times n$ à coefficients dans l'anneau euclidien A, non-nulle. Alors on peut faire des opérations sur les lignes et les colonnes de M qui mettent cette matrice sous la forme*

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

avec d_i qui divise d_{i+1}. De plus, il existe un algorithme explicite pour faire ceci.

On dit que la matrice obtenue est la *forme normale de Smith* de M (ce n'est pas complètement standard, sauf peut-être pour $A = \mathbb{Z}$). On peut montrer qu'elle est unique, mais nous ne le ferons pas.

Démonstration. On procède par récurrence sur n, et le point délicat est de montrer que l'on peut faire des opérations sur M afin de la mettre sous la forme

$$\begin{pmatrix} d_1 & 0 \\ 0 & M' \end{pmatrix} \qquad (*)$$

où M' est une matrice $(n-1) \times (n-1)$ avec la propriété que d_1 divise tous les coefficients de M' dans l'anneau A. En effet, d_1 continuera de diviser les coefficients de M' même si on fait des opérations sur celle-ci, et si on suppose

par récurrence que l'on peut mettre M' sous la forme

$$\begin{pmatrix} d_2 & 0 & \cdots & 0 \\ 0 & d_3 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

avec d_i qui divise d_{i+1} pour $i \geq 2$, alors ceci met bien M sous la forme annoncée.

Cherchons donc à mettre M sous la forme (*). Dans la suite, nous écrivons (a_{ij}) pour les coefficients de M *à l'instant où on parle.*

1. On permute les lignes et les colonnes de M jusqu'à ce que a_{11} soit l'élément de la matrice avec $v(a_{11})$ minimal.
2. On écrit la division euclidienne de a_{12} par a_{11} :

   $$a_{12} = qa_{11} + r$$

   avec $r = 0$ ou $v(r) < v(a_{11})$. Ensuite, on fait $C_2 \leftarrow C_2 - qC_1$, c'est-à-dire qu'on enlève q fois la colonne 1 à la colonne 2, de sorte qu'on a r sur la ligne 1, dans la colonne 2. Si $r \neq 0$, on retourne à l'étape 1 (!!), ce qui aura en particulier pour effet de mettre r en position $(1, 1)$; autrement dit a_{11} devient r dans ce cas.

   On continue jusqu'à avoir un reste nul. On est certain d'y arriver, puisque la suite des entiers $v(a_{11}) \in \mathbb{N}$ que l'on obtient est strictement décroissante.

   Enfin, on recommence avec tous les coefficients de la ligne 1 qui sont non-nuls, puis avec tous ceux de la colonne 1 qui sont non-nuls (en faisant des opérations sur les lignes, cette fois).
3. À ce stade, nous avons mis M sous la forme

   $$\begin{pmatrix} a_{11} & 0 \\ 0 & M' \end{pmatrix}.$$

   Si a_{11} divise tous les coefficients de M', nous avons terminé. Sinon, supposons que la division euclidienne de a_{ij} par a_{11} donne un reste r non-nul. On ajoute alors la colonne j (celle qui contient a_{ij}) à la première

colonne, ce qui fait apparaître r en position $(i, 1)$. Ensuite, on retourne à l'étape 2 (!!!), et en particulier $v(a_{11})$ va décroître strictement (puisque $v(r) < v(a_{11})$). Pour les mêmes raisons que précédemment, ce problème ne peut apparaître qu'un nombre fini de fois, et on va finir par avoir tous les coefficients de M' divisible par a_{11}.

□

Corollaire 116 (Théorème de la base adaptée)**.** *Soit A un anneau euclidien, soit $n \geq 1$ un entier, et soit K un sous-module de A^n. Alors il existe un entier $m \leq n$ tel que K est libre de rang m. Plus précisément, il existe une base $e_1, \ldots, e_n$ de A^n et des éléments $d_1, \ldots, d_m \in A$ non nuls, avec d_i qui divise d_{i+1}, tels que $d_1 e_1, \ldots, d_m e_m$ est une base de K.*

Démonstration. Nous avons vu que, l'anneau A étant euclidien, le sous-module K est de type fini, et en fait il peut être engendré par m_0 éléments avec $m_0 \leq n$. Soit donc $f_0 : A^{m_0} \longrightarrow A^n$ tel que $\mathrm{Im}(f_0) = K$. Nous allons tout de suite remplacer f_0 par $f : A^n \longrightarrow A^n$ définie par $f(x_1, \ldots, x_n) = f_0(x_1, \ldots, x_{m_0})$. Par construction, $\mathrm{Im}(f) = \mathrm{Im}(f_0)$, mais nous pouvons appliquer le théorème précédent à f.

En effet l'application $f : A^n \longrightarrow A^n$ est donnée par sa matrice M dans la base canonique – exactement comme dans la situation sur un corps, que vous connaissez bien depuis le L1 (on a donc $f(v) = Mv$, en voyant v comme un vecteur-colonne à coefficients dans A).

D'après le théorème, il existe des matrices inversibles P et Q, à coefficients dans A, de taille $n \times n$, telles que

$$PMQ = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & d_n \end{pmatrix} = D$$

avec d_i qui divise d_{i+1}. Soit alors $\varepsilon_i \in A^n$ la i-ème colonne de Q, puis $e_i \in A^n$

la i-ème colonne de P^{-1}, et enfin

$$b_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in A^n$$

avec le 1 en i-ième position. Alors $\varepsilon_i = Qb_i$ et $e_i = P^{-1}b_i$. Bien sûr la famille $b_1, \ldots, b_n$ est une base (c'est la base canonique !) de A^n, et puisque Q et P^{-1} sont inversibles, on en déduit que $\varepsilon_1, \ldots, \varepsilon_n$ est une base, ainsi que $e_1, \ldots, e_n$.

Par ailleurs on a

$$f(\varepsilon_i) = MQb_i = P^{-1}Db_i = P^{-1}d_ib_i = d_ie_i\,.$$

On en déduit que $\mathrm{Im}(f)$ est engendré par les éléments $d_1e_1, \ldots, d_ne_n$. Maintenant, soit m le plus grand entier tel que $d_i \neq 0$; on note que si $d_i = 0$ alors $d_{i+1} = 0$ puisque d_i divise d_{i+1}, donc les éléments $d_1, \ldots, d_m$ sont tous non-nuls. La famille $d_1e_1, \ldots, d_me_m$ est donc libre, et c'est bien une base de $\mathrm{Im}(f)$. □

§3.4 La classification des modules de type fini

Théorème 117. *Soit A un anneau euclidien, et soit V un module de type fini. Alors il existe un isomorphisme*

$$V \cong A^r \times A/(d_1) \times A/(d_2) \times \cdots \times A/(d_k)$$

où les éléments $d_i \in A$ ne sont pas des unités, et ne sont pas nuls, et vérifient $d_i | d_{i+1}$.

Démonstration. Les arguments ont été en grande partie donnés au cours du chapitre ; reprenons. Par définition il existe une application surjective

$$\pi : A^n \longrightarrow V$$

dont le noyau sera noté K. On a donc $V \cong A^n/K$.

Par le théorème de la base adaptée, il existe une base $e_1, \dots, e_n$ de A^n et des éléments $d_1, \dots, d_m \in A$ non nuls, avec d_i qui divise d_{i+1}, tels que $d_1 e_1, \dots, d_m e_m$ est une base de K.

Nous aurons besoin de

$$\varphi : A^n \longrightarrow A^n$$

$$\varphi(x_1, \dots, x_n) = x_1 e_1 + \cdots + x_n e_n \,.$$

La composée $\pi \circ \varphi : A^n \longrightarrow V$ est surjective, et son noyau est $\varphi^{-1}(K)$; en clair le noyau N est composé des éléments de la forme

$$(a_1 d_1, \dots, a_m d_m, 0, \dots, 0)$$

avec $a_i \in A$. On a donc $V \cong A^n/N$.

Mais N est aussi le noyau de l'application évidente

$$A^n \longrightarrow A/(d_1) \times A/(d_2) \times \cdots \times A/(d_m) \times A^{n-m} \,,$$

donc A^n/N est bien isomorphe au module qui apparaît à droite ici, d'où le résultat avec $r = n - m$.

Notons que l'on peut effectivement supposer que d_i n'est pas une unité, car dans le cas contraire on a $(d_i) = A$ et $A/(d_i) = \{0\}$, donc on peut tout simplement retirer d_i de la liste. □

Il existe aussi un théorème d'unicité :

Théorème 118. *Supposons qu'il existe un isomorphisme*

$$A^r \times A/(d_1) \times \cdots \times A/(d_k) \cong A^{r'} \times A/(d'_1) \times \cdots \times A/(d'_{k'})$$

où les d_i et les d'_i sont non-nuls et non-inversibles, avec $d_i | d_{i+1}$ et $d'_i | d'_{i+1}$. Alors $r' = r$, $k' = k$, et après renumérotation, on a $d'_i = u_i d_i$ avec $u_i \in A^\times$.

Nous ne montrerons pas ce théorème au tableau (c'est vraiment trop long) ; voir l'appendice à la fin du chapitre.

Les éléments d_i sont donc bien définis, à un inversible près, et on les appelle les *diviseurs élémentaires* du module.

On va reformuler ce théorème de classification encore et encore, d'abord dans le cas général, puis en spécialisant à $A = \mathbb{Z}$ puis $A = \mathbb{K}[X]$. Commençons par :

Théorème 119 (Deuxième théorème de classification)**.** *Soit A un anneau euclidien, et soit V un module de type fini. Alors il existe un isomorphisme*

$$V \cong A^r \times P$$

où P est un produit de modules de la forme

$$A/(p^\alpha)\,,$$

avec $p \in A$ un élément premier, et $\alpha \geq 1$. Cette décomposition est unique.

Démonstration. C'est essentiellement le lemme chinois. Supposons que l'on ait écrit l'élément $d \in A$ comme produit de premiers :

$$d = up_1^{\alpha_1} \cdots p_s^{\alpha_s}\,,$$

avec $u \in A^\times$, chaque p_i premier, et pour $i \neq j$ les éléments p_i et p_j premiers entre eux. Alors $a = p_1^{\alpha_1}$ et $b = p_2^{\alpha_2} \cdots p_s^{\alpha_s}$ sont premiers entre eux. Si $I = (a)$ et $J = (b)$, alors $I + J = A$ par Bézout, et $(d) = (p_1^{\alpha_1} \cdots p_s^{\alpha_s}) = IJ$. Par le lemme chinois, on a

$$A/(d) \cong A/(p_1^{\alpha_1}) \times A/(p_2^{\alpha_2} \cdots p_s^{\alpha_s})\,.$$

En recommençant un certain nombre de fois, on finit par avoir

$$A/(d) \cong A/(p_1^{\alpha_1}) \times \cdots \times A/(p_s^{\alpha_s})$$

Il suffit alors d'appliquer ça à chaque terme $A/(d_i)$ du théorème de classification. Attention, un nombre premier donné va apparaître plusieurs fois (s'il apparaît dans d_1, il apparaît aussi dans d_2, d_3, etc). □

Il est probablement utile d'énoncer le théorème de classification dans le cas $A = \mathbb{Z}$:

Théorème 120 (Classification des groupes abéliens de type fini)**.** *Soit V un groupe abélien de type fini. Alors il existe un isomorphisme*

$$V \cong \mathbb{Z}^r \times \mathbb{Z}/d_1 \times \cdots \times \mathbb{Z}/d_k$$

avec $d_i \geq 2$ et $d_i \mid d_{i+1}$. Il existe aussi un isomorphisme

$$V \cong \mathbb{Z}^r \times P$$

où P est un produit direct de groupes abéliens de la forme

$$\mathbb{Z}/(p^\alpha)$$

avec p premier et $\alpha \geq 1$.

Ces décompositions sont uniques.

En particulier, si V est un groupe abélien fini, alors il admet des décompositions comme ci-dessus avec $r = 0$.

Exemple 121. Prenons $A = \mathbb{Z}$ et

$$V = \mathbb{Z}/15 \times \mathbb{Z}/24\,.$$

(On va écrire $\mathbb{Z}/n$ au lieu de $\mathbb{Z}/(n)$ ou $\mathbb{Z}/n\mathbb{Z}$.) L'écriture ci-dessus n'est pas celle des théorèmes de classification. Mais le lemme chinois nous dit que

$$\mathbb{Z}/15 \cong \mathbb{Z}/3 \times \mathbb{Z}/5$$

et

$$\mathbb{Z}/24 \cong \mathbb{Z}/3 \times \mathbb{Z}/8\,.$$

Ainsi

$$A \cong \mathbb{Z}/(2^3) \times \mathbb{Z}/3 \times \mathbb{Z}/3 \times \mathbb{Z}/5\,.$$

C'est la forme du deuxième théorème.

Pour obtenir la forme du premier théorème, on écrit

$$\mathbb{Z}/(2^3) \times \mathbb{Z}/3 \times \mathbb{Z}/5 \cong \mathbb{Z}/120\,,$$

d'où

$$A \cong \mathbb{Z}/3 \times \mathbb{Z}/120\,.$$

Puisque 3 divise 120, les diviseurs élémentaires de V sont bien 3 et 120.

Dans cet exemple, nous avons joué avec le lemme chinois pour passer d'un théorème de classification à l'autre. Il est très facile de se convaincre qu'on peut toujours y arriver, et ceci ne doit pas masquer ce qui est vraiment non-trivial dans les théorèmes ci-dessus : à savoir, qu'un groupe abélien de type fini V est toujours un produit de groupes cycliques !

§3.5 Application à la reduction des endomorphismes

On va maintenant explorer le théorème de classification dans le cas de $A = \mathbb{K}[X]$. Commençons par un exemple.

Exemple 122. Prenons $\mathbb{K} = \mathbb{Q}$, et soit $V = \mathbb{Q}^3$ muni de l'endomorphisme f dont la matrice est

$$F = \begin{pmatrix} -2 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Que nous dit la théorie développé jusqu'ici du $\mathbb{Q}[X]$-module (V, f) ? Dans ce premier exemple, nous supposerons que nous disposons d'un ordinateur, ou alors que nous n'avons pas peur de passer du temps sur les calculs. Plutôt que d'appliquer un théorème de classification, nous reprenons le raisonnement, en le rendant explicite au passage.

L'argument de l'exemple 114 nous dit que, en posant $A = \mathbb{Q}[X]$, l'application $\varphi : A^3 \longrightarrow A^3$ dont la matrice est $M = F - XI$ vérifie que $A^3/\mathrm{Im}(\varphi) \cong V$. Or si nous appliquons l'algorithme qui donne la forme de Smith de M, nous trouvons (après pas mal d'étapes !) une matrice diagonale avec les coefficients $(1, X, X^2)$. On en déduit, comme dans la preuve des théorèmes de classification, que

$$V \cong A^3/\mathrm{Im}(\varphi) \cong A/(1) \times A/(X) \times A/(X^2).$$

Ici $(1) = A$ donc $A/(1)$ est nul ; on peut donc oublier ce facteur. Par ailleurs, on sait depuis un exercice de la feuille décrire $\mathbb{K}[X]/(D)$: il existe une base dans laquelle la multiplication par X est donnée par la matrice compagne de D. Dans l'exemple, la matrice compagne de $D = X$ est la matrice nulle de taille 1×1 ; la matrice compagne de $D = X^2$ est

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

La conclusion est qu'il existe une base de V dans laquelle l'endomorphisme

f a pour matrice

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Pour un endomorphisme qui n'est pas diagonalisable, c'est plutôt bien : nous avons trouvé une matrice pour f avec beaucoup de zéros !

En pratique, on ne procèdera presque jamais comme dans l'exemple précédent (qui était absent de la première version de ce poly, d'ailleurs). Le problème est que rapidement, le calcul de la forme de Smith devient énorme. Avec un ordinateur, ce n'est pas un réel souci, mais dans ce cours on veut se concentrer sur les calculs faisables « à la main ». Par ailleurs, on va aussi rencontrer des questions plus théoriques, du genre « comment démontrer le théorème de Cayley-Hamilton avec la théorie des $\mathbb{K}[X]$-modules ? », et bien d'autres, qui sont plus intéressantes que les questions du type « mettre cette matrice sous une forme avec beaucoup de zéros ».

On va donc faire autrement : on énonce un théorème de classification spécialisé à $\mathbb{K}[X]$ le plus précis possible, et on essaie de toujours s'y ramener. Voir l'exemple 127 ci-dessous pour une autre version du calcul que l'on vient de faire, qui est la façon « standard » de procéder.

Théorème 123. *Soit $\mathbb{K}$ un corps commutatif et V un espace vectoriel de dimension finie (et > 0) sur $\mathbb{K}$. Soit $f : V \longrightarrow V$ un endomorphisme.*

Alors le $\mathbb{K}[X]$-module (V, f) est isomorphe à un produit de modules de la forme

$$V \cong \mathbb{K}[X]/(P_1^{\alpha_1}) \times \cdots \times \mathbb{K}[X]/(P_s^{\alpha_s})$$

où P_i est un polynôme irréductible et unitaire. Une telle décomposition est unique.

De plus, (V, f) est indécomposable si et seulement si $s = 1$; il est simple si et seulement si $s = 1$ et $\alpha_1 = 1$.

En termes de matrices, ceci signifie qu'il existe une base de V dans laquelle

la matrice de f est de la forme (diagonale par blocks) ci-dessous :

$$\begin{pmatrix} Co(P_1^{\alpha_1}) & & & \\ & Co(P_2^{\alpha_2}) & & \\ & & \ddots & \\ & & & Co(P_s^{\alpha_s}) \end{pmatrix}$$

où on écrit $Co(P)$ pour la matrice compagne du polynôme P.

Le polynôme caractéristique de f est $P_1^{\alpha_1} \cdots P_s^{\alpha_s}$, et le polynôme minimal de f est le ppcm des polynômes $P_i^{\alpha_i}$.

Démonstration. L'existence et l'unicité de la décomposition de V proviennent directement du deuxième théorème de classification. On note qu'il n'y a pas de facteur $\mathbb{K}[X]^r$ avec $r > 0$ puisque, dans le cas contraire, V serait de dimension infinie. On note également que l'on peut prendre chaque P_i unitaire car $(P_i^{\alpha_i}) = (\lambda P_i^{\alpha_i})$ si $\lambda \in \mathbb{K}$ est une constante.

Si V indécomposable, alors il ne peut y avoir qu'un facteur dans la décomposition, donc $s = 1$. Réciproquement, si $V = U_1 \oplus U_2$ avec $U_i \neq \{0\}$, alors en appliquant la décomposition à U_1 et U_2, on voit que $s > 1$.

Si V est simple, alors il est indécomposable, donc $s = 1$. Le module $\mathbb{K}[X]/(P_1^{\alpha_1})$ est simple si et seulement si $P_1^{\alpha_1}$ est premier, d'après la proposition 109, donc si et seulement si $\alpha_1 = 1$. □

Les résultats classiques de « réduction des endomorphismes » se retrouvent tous très facilement. Par exemple :

Corollaire 124 (Cayley-Hamilton)**.** *Le polynôme minimal divise le polynôme caractéristique.*

Démonstration. Évident d'après la description de ces polynômes dans le théorème. □

La tradition est de reformuler le théorème de classification pour que ça deviennent *la théorie de Jordan*. Il s'agit d'une série de remarques simples pour amener les matrices sous des formes agréables.

Définition 125. Soit $\lambda \in \mathbb{K}$ et $n \geq 1$ un entier. On note

$$J_n(\lambda) = \begin{pmatrix} \lambda & 0 & \cdots & & 0 \\ 1 & \ddots & \ddots & & \vdots \\ 0 & 1 & \ddots & \ddots & \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}.$$

ou en d'autres termes, $J_n(\lambda) = \lambda I + Co(X^n)$. On appelle $J_n(\lambda)$ le *bloc de Jordan de taille n associé à la valeur propre* λ.

En particulier $J_n(0) = Co(X^n)$, et il est important d'observer les puissances de cette matrice :

$$J_n(0)^2 = \begin{pmatrix} 0 & 0 & \cdots & & 0 \\ 0 & \ddots & \ddots & & \vdots \\ 1 & 0 & \ddots & \ddots & \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 1 & 0 & 0 \end{pmatrix},$$

les 1 « descendent », puis on arrive à

$$J_n(0)^{n-1} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 1 & \cdots & 0 \end{pmatrix},$$

et enfin $J_n(0)^n = 0$. Par suite, on a $(J_n(\lambda) - \lambda I)^n = 0$.

Lemme 126. *Soit $f : V \longrightarrow V$ un endomorphisme nilpotent de l'espace vectoriel V de dimension finie sur $\mathbb{K}$. Alors il existe une base de V dans laquelle la matrice de f est diagonale par blocs, de la forme*

$$\begin{pmatrix} J_{\alpha_1}(0) & & & \\ & J_{\alpha_2}(0) & & \\ & & \ddots & \\ & & & J_{\alpha_k}(0) \end{pmatrix}.$$

De plus, les entiers $\alpha_1, \ldots, \alpha_k$ sont uniques, à renumérotation près.

Démonstration. D'après le théorème de classification, (V, f) est isomorphe à un produit de modules de la forme $\mathbb{K}[X]/(P_i^{\alpha_i})$ avec P_i irréductible. Le ppcm des $P_i^{\alpha_i}$ est le polynôme minimal ; or par hypothèse, il y a un polynôme de la forme X^m qui annule f, donc le polynôme minimal, qui le divise, est de la forme X^k, donc $P_i = X$ pour tout i. Puisque $Co(X^{\alpha_i}) = J_{\alpha_i}(0)$, le théorème de classification donne bien la forme ci-dessus. L'unicité est une conséquence de l'énoncé d'unicité dans le théorème de classification. □

Exemple 127. Arrêtons nous pour voir un exercice absolument typique : on se donne par exemple

$$N = \begin{pmatrix} -2 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

et la question est « montrer que N est nilpotente et donner sa forme de Jordan ». (C'est la même matrice que dans l'exemple 122, et donc comme promis, on attaque le même calcul différemment). On calcule $N^2 = 0$ donc N est bien nilpotente. Pour la forme de Jordan, on a *a priori* trois possibilités :

$$\begin{pmatrix} J_1(0) & & \\ & J_1(0) & \\ & & J_1(0) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

ou bien

$$\begin{pmatrix} J_1(0) & \\ & J_2(0) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

ou encore

$$J_3(0) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Or $N \neq 0$ donc le premier cas est exclu, et $N^2 = 0$ donc le troisième cas est exclu. La forme de Jordan, par élimination, est donc celle du milieu.

Cet argument est très facile évidemment, mais vous voyez bien qu'avec une matrice disons 10×10, le nombre de formes de Jordan potentielles aurait été très grand, et ça se complique. Pour simplifier un peu, il est bon d'énoncer :

Lemme 128. *Le nombre de blocs de Jordan, pour un endomorphisme nilpotent, est la dimension du noyau.*

Démonstration. On peut trouver une base dans laquelle la matrice de l'endomorphisme f est formée de k blocs de Jordan. Pour chaque bloc, il y a une colonne nulle, donc le noyau de f est certainement de dimension $\geq k$; mais les autres colonnes sont linéairement indépendantes, donc le rang de f est $\geq n-k$, ce qui impose que la dimension du noyau est $\leq k$ (théorème du rang). Finalement $\dim \ker(f) = k$. □

Dans l'exemple précédent, la matrice N est visiblement de rang 1, donc la forme de Jordan comporte deux blocs. Et c'est fini !

Notons aussi qu'il y a un exercice de la feuille (numéroté 60 en 2022-2023) qui approfondit l'idée de ce lemme, et ramène tous les calculs à un système assez simple.

Pour les endomorphismes nilpotents, on a appliqué le théorème de classification de manière très directe. C'est pour le lemme suivant qu'il y a une petite astuce de calcul :

Lemme 129. *Soit $V = \mathbb{K}[X]/(X-\lambda)^n$, vu comme $\mathbb{K}[X]$-module avec l'endomorphisme f qui multiplie par X. Alors il existe une base de V dans laquelle f a pour matrice le bloc de Jordan $J_n(\lambda)$.*

Démonstration. On applique le lemme précédent à $g = f - \lambda I$, qui vérifie $g^n = 0$ puisque $(f-\lambda I)^n = 0$. Il existe donc une base de V dans laquelle g a une matrice diagonale par blocs, avec des blocs de Jordan $J_{\alpha_1}(0), \ldots, J_{\alpha_k}(0)$.

On fait alors la remarque suivante : un sous-espace vectoriel $U \subset V$ est stable par f si et seulement s'il est stable par g (étant donné qu'il est toujours stable par λI !). Or V est indécomposable pour l'endomorphisme f, donc il est également indécomposable pour g. Donc il n'y a qu'un seul bloc de Jordan, et g a pour matrice $J_n(0)$ dans la base ci-dessus. Ainsi $f = \lambda I + g$ a pour matrice $\lambda I + J_n(0) = J_n(\lambda)$. □

Théorème 130 (Forme de Jordan)**.** *Soit V un espace vectoriel de dimension finie sur le corps commutatif $\mathbb{K}$, et soit $f : V \longrightarrow V$ un endomorphisme. On suppose que le polynôme caractéristique de f est scindé (ce qui est automatique si $\mathbb{K} = \mathbb{C}$, par exemple), disons*

$$\chi_f = \prod_{i=1}^{r} (X - \lambda_i)^{m_i} .$$

Alors il existe une base de V dans laquelle f a une matrice diagonale par blocs, avec des blocs de Jordan. Cette forme est unique et on l'appelle la forme de Jordan *de f.*

Plus précisément, pour chaque i il existe une base de $\ker(f - \lambda_i I)^{m_i}$ *dans laquelle f a une matrice de la forme*

$$\begin{pmatrix} J_{\alpha_{i,1}}(\lambda_i) & & & \\ & J_{\alpha_{i,2}}(\lambda_i) & & \\ & & \ddots & \\ & & & J_{\alpha_{i,k_i}}(\lambda_i) \end{pmatrix} .$$

De plus,

$$V = \bigoplus_{i=1}^{k} \ker(f - \lambda_i I)^{m_i} .$$

En particulier, on constate que $\ker(f - \lambda_i I)^{m_i}$ *est de dimension m_i.*

Démonstration. D'après le théorème de classification, V est isomorphe à un produit de modules de la forme $\mathbb{K}[X]/(P^\alpha)$, avec P irréductible. Le produit de tous les P^α est χ_f, donc chaque P de la forme $X - \lambda_i$ pour un certain i. D'après le dernier lemme, on peut trouver une base de $\mathbb{K}[X]/(P^\alpha) = \mathbb{K}[X]/(X - \lambda_i)^\alpha$ dans laquelle l'endomorphisme de multiplication par X a pour matrice $J_\alpha(\lambda_i)$. Ainsi, il existe une base de V dans laquelle f a une matrice diagonale par blocs, avec tous les $J_\alpha(\lambda_i)$ sur la diagonale.

Il faut plus de notations (malheureusement) pour obtenir l'énoncé plus précis. Soit V_i le produit de tous les facteurs de la forme $\mathbb{K}[X]/(X - \lambda_i)^{\alpha_{i,j}}$ donnés par le théorème, de sorte que

$$V_i = \mathbb{K}[X]/(X - \lambda_i)^{\alpha_{i,1}} \times \mathbb{K}[X]/(X - \lambda_i)^{\alpha_{i,2}} \times \cdots \times \mathbb{K}[X]/(X - \lambda_i)^{\alpha_{i,k_i}}$$

(avec $\alpha_{i,1} + \cdots + \alpha_{i,k_i} = m_i$) et on sait donc qu'il existe un isomorphisme de $\mathbb{K}[X]$-modules

$$\varphi : V_1 \oplus V_2 \oplus \cdots \oplus V_r \longrightarrow V .$$

La matrice de l'endomorphisme de multiplication par X sur V_i est exactement celle donnée dans le théorème. Il nous reste à montrer que $\varphi(V_i) = \ker(f - \lambda_i I)^{m_i}$.

Or φ est un isomorphisme de $\mathbb{K}[X]$-modules, donc la question revient à montrer que $V_i = K_i$, en posant

$$K_i = \{v \in V_1 \oplus \cdots \oplus V_r \mid (X - \lambda_i I)^{m_i} \cdot v = 0\} .$$

D'un côté, si nous regardons $\mathbb{K}[X]/(X-\lambda_i)^{\alpha_{i,j}}$, nous savons que la multiplication par $(X-\lambda_i)^{\alpha_{i,j}}$ est nulle sur ce module, et donc la multiplication par $(X-\lambda_i)^{m_i}$ est également nulle, puisque $\alpha_{i,j} \leq m_i$; on en déduit que $\mathbb{K}[X]/(X-\lambda_i)^{\alpha_{i,j}} \subset K_i$. C'est vrai pour tout j, donc $V_i \subset K_i$.

Pour l'inclusion réciproque, prenons $v = v_1 + \cdots + v_r$ avec chaque $v_j \in V_j$. La somme étant directe, si $(X - \lambda_i)^{m_i} \cdot v = 0$, c'est que pour tout j on a $(X - \lambda_i)^{m_i} \cdot v_j = 0$. Or v_j est annulé par $(X - \lambda_j)^{m_j}$, et si $i \neq j$, alors par Bézout il existe des polynômes A et B tels que

$$A(X - \lambda_i)^{m_i} + B(X - \lambda_j)^{m_j} = 1 .$$

On applique ceci à v_j, des deux côtés, et on en déduit $v_j = 0$ pour $i \neq j$. Il reste $v = v_i \in V_i$, donc $K_i \subset V_i$.

La dimension de V_i étant évidemment m_i, on en déduit la dernière phrase du théorème. □

Nous allons voir de nombreux exemples dans les exercices. Indiquons tout de suite, cependant :

Lemme 131. *Le nombre de blocs de Jordan associés à la valeur propre λ_i est la dimension de* $\ker(f - \lambda_i)$.

Démonstration. Comme pour un très grand nombre de choses concernant les formes de Jordan, on démontre ceci en se ramenant au cas nilpotent : on considère la restriction de $f - \lambda_i I$ au sous-espace $\ker(f - \lambda_i I)^{m_i}$, et on applique le lemme 128. □

§3.6 Appendice : démonstration de l'unicité

Soit A un anneau euclidien. Nous allons montrer l'unicité de la décomposition des A-modules donnée par les théorèmes de classification. Ici nous allons travailler avec la deuxième forme ; c'est une exercice sur le lemme chinois, que l'on vous laisse, que d'en déduire l'unicité avec la première forme.

Concrètement, on se donne un entier $r \geq 0$; des premiers $p_1, p_2, \ldots, p_k \in A$ avec $(p_i) \neq (p_j)$ si $i \neq j$; pour chaque i on se donne des entiers $1 \leq \alpha_{i1} \leq \alpha_{i2} \leq \cdots \leq \alpha_{im_i}$; et on pose

$$V = A^r \times \prod_{i,j} A/(p_i^{\alpha_{ij}}) \,.$$

De manière analogue, soit

$$V' = A^s \times \prod_{i,j} A/(q_i^{\beta_{ij}}) \,,$$

pour des premiers $q_1, \ldots, q_\ell$. On suppose que $V \cong V'$, et on va montrer que $r = s$, que $k = \ell$, puis que, quitte à renuméroter, on a $(p_i) = (q_i)$ pour tout i, et enfin que pour tout i on a les mêmes puissances de p_i, c'est-à-dire que $m_i = m'_i$ et $\alpha_{ij} = \beta_{ij}$ pour tous les i, j.

Lorsque M est un A-module, on définit le *sous-module de torsion* M_{tors} par

$$M_{tors} = \{m \in M \mid \exists a \in A, a \neq 0, a \cdot m = 0\} \,.$$

On vérifie sans peine que si on a un isomorphisme $M \cong N$, alors il induit des isomorphismes $M_{tors} \cong N_{tors}$, mais aussi $M/M_{tors} \cong N/N_{tors}$.

Appliquons ceci à V et V' définis comme ci-dessus. Nous avons

$$V_{tors} = \{0\} \times \prod_{i,j} A/(p_i^{\alpha_{ij}}) \quad \text{et} \quad V/V_{tors} \cong A^r \,,$$

et des résultats similaires pour V'. En particulier, il y a un isomorphisme $A^r \cong A^s$. Commençons donc par montrer :

Lemme 132. *Supposons qu'il existe un isomorphisme $A^r \cong A^s$. Alors $r = s$. (Ceci pour A euclidien, mais la preuve va marcher pour A principal, voire mieux.)*

Démonstration. Si A est un corps, posons $I = (0)$. Sinon, il existe un élément premier $p \in A$, et on pose $I = (p)$. Dans tous les cas, l'anneau A/I est un corps. (Dans un anneau commutatif quelconque, il faudrait juste montrer qu'il existe un idéal maximal I pour que la démonstration fonctionne, et on peut le faire avec le lemme de Zorn.)

Un isomorphisme $V \cong V'$ de A-modules induit un isomorphisme $V/IV \cong V'/IV'$ de (A/I)-modules. Ici, on obtient $(A/p)^r \cong (A/p)^s$ comme espaces vectoriels sur le corps A/p. En comparant les dimensions, on trouve $r = s$. □

Quitte à remplacer V et V' par V_{tors} et V'_{tors} respectivement, nous pouvons continuer la démonstration en supposant $r = s = 0$.

Lemme 133. *Soient p, q des premiers de A, soit $M = A/(p^\alpha)$, $I = (q^\beta)$, et $J = (q^\gamma)$ où α, β et γ sont des entiers. On suppose que $\gamma \leq \beta$ de sorte que $I \subset J$. Alors :*

- *Si $(q) \neq (p)$, on a $IM = JM = M$, d'où $M/IM = \{0\}$ et $JM/IM = \{0\}$.*
- *Si $(q) = (p)$ et $\beta \geq \alpha$, on a $IM = \{0\}$ d'où $M/IM = M$ et $JM/IM = JM$.*
- *Si $(q) = (p)$ et $\beta \leq \alpha$, on a $M/IM \cong A/I$ et $JM/IM \cong A/(p^{\beta-\gamma})$.*

(Noter que pour $\alpha = \beta$ les deux derniers points donnent bien le même résultat.)

En particulier, pour tout $\alpha \geq 1$, on a $M/(p)M \cong A/(p)$ (c'est le cas $\beta = 1$, $(p) = (q)$).

Autre cas particulier (obtenu pour $\gamma = \beta - 1$), on note que $(p^{\beta-1})M/(p^\beta)M$ est isomorphe à $A/(p)$ lorsque $(p) = (q)$ et $\beta \leq \alpha$, alors qu'il est nul dans tous les autres cas.

Démonstration. Si $(p) \neq (q)$, alors $(p^\alpha, q^\beta) = A$ (un pgcd de p^α et q^β doit les diviser tous les deux, donc c'est une unité). On peut donc écrire $p^\alpha u + q^\beta v = 1$ pour des $u, v \in A$, et pour $m \in M$ on en déduit

$$m = p^\alpha um + q^\beta vm = q^\beta vm \in IM\,.$$

D'où $IM = M$ et $M/IM = \{0\}$. De même $JM = M$. On continue en supposant $p = q$.

Si $\beta \geq \alpha$, il est clair que $p^\beta m = 0$ pour tout $m \in M$, donc $IM = \{0\}$.

Reste le cas où $\beta \leq \alpha$. On considère la composition $A \longrightarrow M \longrightarrow M/IM$ des applications de quotient, et on constate qu'elle est évidemment surjective de noyau I, d'où le premier point. Pour le deuxième, on considère l'homomorphisme $A \longrightarrow JM$ définie par $a \mapsto p^\gamma \overline{a}$, où $\overline{a}$ désigne la classe de $a \in A$ dans $M = A/(p^\alpha)$; il est facile de voir qu'il est surjectif. Si on considère l'application $A \longrightarrow JM/IM$ obtenue en composant avec $JM \longrightarrow JM/IM$, le noyau est visiblement $(p^{\beta-\gamma})$. □

Appliquons ceci à notre module V (avec la remarque que si $M = M_1 \times M_2$, alors $M/IM \cong M_1/IM_1 \times M_2/IM_2$, bien sûr). On prend un premier $p \in A$, on pose $I = (p)$, et on regarde V/IV : d'après le lemme et la remarque, on obtient $V/IV = 0$ si aucun des p_i ne vérifie $(p_i) = (p)$, alors que s'il existe un indice i tel que $(p_i) = (p)$, on obtient $V/IV \cong (A/p)^{m_i}$. C'est un isomorphisme de A-modules mais aussi d'espaces vectoriels sur le corps $A/(p_i)$.

Puisqu'on suppose que $V \cong V'$, on constate en faisant le même raisonnement avec V' et en considérant l'isomorphisme $V/IV \cong V'/IV'$ que, si p_i est l'un des premiers qui apparaissent dans la définition de V, alors il doit y avoir un q_j tel que $(q_j) = (p_i)$ (sinon V'/IV' serait nul et V/IV non nul). Quitte à renuméroter, on peut supposer que $(p_i) = (q_i)$ pour tout i. De plus, on a $m_i = m'_i$, en regardant les dimensions dans l'isomorphisme ci-dessus.

Il reste à montrer que les nombres α_{ij} et les nombres β_{ij} coïncident. Plus précisément, fixons i, et considérons les nombres $\alpha_{i1} \leq \alpha_{i2} \leq \cdots \leq \alpha_{im_i}$ d'une part, et les nombres $\beta_{i1} \leq \beta_{i2} \leq \cdots \leq \beta_{im_i}$ d'autre part ; on souhaite montrer que $\alpha_{ij} = \beta_{ij}$ pour tout j. Notons $p = p_i$. Pour tout entier n on définit

$$d_n = \dim(p^{n-1})V/(p^n)V = \dim(p^{n-1})V'/(p^n)V' \,,$$

où il s'agit des dimensions de ces espaces vectoriels sur le corps $A/(p)$. D'après le cas particulier mis en évidence juste après le dernier lemme, on a

$$d_n = \text{le nombre de } j \text{ tels que } \alpha_{ij} \geq n \,.$$

Bien sûr, on a une description similaire avec β_{ij} qui remplace α_{ij}. La démonstration sera donc entièrement terminée par le lemme suivant :

Lemme 134. *Soit $x_1 \leq x_2 \leq \cdots \leq x_m$ une suite finie d'entiers, et soit d_n le nombre d'indices j tels que $x_j \geq n$, pour tout $n \in \mathbb{N}$. Alors les nombres d_n déterminent la suite $x_1, \ldots, x_m$.*

Démonstration. Pour être précis, on suppose que l'on connaît m et tous les nombres d_n pour $n \in \mathbb{N}$, et on doit montrer que l'on peut reconstituer $x_1, \ldots, x_m$. On procède par récurrence sur m. Pour $m = 1$ il n'y a qu'un nombre x_1, et c'est le plus grand entier tel que $d_{x_1} \neq 0$, donc il est entièrement déterminé.

Voyons l'hérédité. On commence de la même manière : on a $d_n = 0$ pour $n > x_m$, et en fait le nombre x_m est le plus grand entier tel que $d_{x_m} \neq 0$. Donc x_m est déterminé par les d_n, ainsi que le nombre de fois que x_m apparait dans la suite, puisque c'est précisément $d = d_{x_m}$.

On a terminé si $d = m$. Sinon, par récurrence, la suite $x_1, \ldots, x_{m-d}$ est déterminée par les nombres d'_n, avec $d'_n =$ le nombre d'indices $j \leq m - d$ tels que $x_j \geq n$; or $d'_n = d_n - d$ pour $n < x_m$, et $d'_n = 0$ sinon, ce qui montre que les d'_n sont déterminés par les d_n. □

Chapitre 4

Groupes

Dans ce chapitre, nous allons aller un peu plus loin avec les groupes, en préparation du chapitre suivant et aussi parce que certains des résultats ci-dessous sont des classiques.

Il faut avoir en tête quelques exemples de groupes, pour apprécier les théorèmes. Nous allons surtout privilégier les groupes finis ici. Il y a donc les groupes abéliens finis, qui – on le sait depuis le chapitre précédent – sont des produits de groupes cycliques. Il y a aussi le groupe diédral D_8, que nous allons examiner dans les exercices : il est engendré par deux éléments r et s, avec $s^2 = r^4 = 1$ et $srs = r^{-1}$, et il compte en tout huit éléments, qui sont $1, r, r^2, r^3$ (ce sont les rotations, dans l'interprétation géométrique que l'on a de D_8) et s, sr, sr^2, sr^3 (qui sont des symétries par rapport à des droites).

Ensuite, nous avons le groupe Q_8. Celui-ci est un sous-groupe de $\mathbb{H}^\times$, où $\mathbb{H}$ est le corps des quaternions, vu dans les exercices ; on définit $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$, il comporte donc 8 éléments, avec $i^2 = j^2 = k^2 = ijk = -1$.

Enfin, dernier exemple important, celui de S_n, le groupe symétrique de degré n, dont les éléments sont les permutations de l'ensemble $\{1, 2, \ldots, n\}$.

§4.1 Les classes à gauche

Lorsque H est un sous-groupe de G, nous allons définir un ensemble G/H qui va être utile à plusieurs titres. Dans certains cas, G/H sera lui-même un groupe, mais nous n'utiliserons pas énormément ces quotients, finalement.

Par contre, on aura besoin de G/H quand nous parlerons des actions de groupes, et également dans un argument classique de comptage qui vient tout de suite.

Premières définitions

Définition 135. Soit G un groupe, et $H \subset G$ un sous-groupe. Pour $g \in G$, on note

$$gH := \{gh \mid h \in H\},$$

et on pose

$$G/H := \{gH \mid g \in G\}.$$

Lorsque G et H sont fixés une fois pour toutes, on note $\overline{g}$ pour gH. Avec cette notation, on a

$$G/H = \{\overline{g} \mid g \in G\}.$$

Jusqu'ici, c'est simplement la notation pour les modules quotients qui a changé ($v + U$ devient gH). Est-ce qu'on peut toujours espérer qu'il y ait une structure de groupe sur G/H, de sorte que l'application $p\colon G \longrightarrow G/H$ définie par $g \mapsto \overline{g}$ soit un homomorphisme ? La réponse est non, pas toujours.

Définition 136. Soit H un sous-groupe de G. On dit que H est *distingué* dans G lorsque, pour tout $g \in G$, on a $gHg^{-1} \subset H$.

Remarque. Dans ce cas, en prenant g^{-1} on a $g^{-1}Hg \subset H$ d'où $H \subset gHg^{-1}$ (en multipliant à gauche par g et à droite par g^{-1}), donc en fait on a une égalité $gHg^{-1} = H$ pour tout $g \in G$.

Exemple 137. Si G est abélien, il est clair que tous ses sous-groupes sont distingués. Si $G = S_n$ (le groupe symétrique sur n symboles), et si $H = \{Id, (1,2)\}$, alors $gHg^{-1} = \{Id, (g(1), g(2))\}$ (nous ferons des révisions sur les permutations plus tard ; c'est juste pour avoir un exemple). Clairement, ce H n'est pas distingué en général (prendre $n \geq 3$ et une permutation g telle que $g(1) = 3$).

Plus important :

Lemme 138. *S'il existe un homomorphisme $\varphi\colon G \longrightarrow G'$ tel que $H = \ker \varphi$, alors H est distingué.*

Démonstration. En effet pour $h \in H$ et $g \in G$

$$\varphi(ghg^{-1}) = \varphi(g)\varphi(h)\varphi(g)^{-1} = \varphi(g)\varphi(g)^{-1} = 1\,,$$

d'où $ghg^{-1} \in H$. □

Conclusion : si H n'est pas distingué, il n'est pas possible que H soit le noyau de quoi que ce soit, donc une structure de groupe sur G/H comme espéré ci-dessus ne peut pas exister.

Mais c'est le seul obstacle :

Lemme 139. *Supposons que H soit un sous-groupe distingué de G. Alors il existe une structure de groupe sur G/H, et une seule, telle que l'application naturelle $G \longrightarrow G/H$ soit un homomorphisme. Son noyau est H.*

Démonstration. Si X et Y sont des parties de G, on définit

$$XY = \{xy \mid x \in X, y \in Y\}\,.$$

Or on peut vérifier que, pour $\sigma, \tau \in G$, on a

$$(\sigma H)(\tau H) = \sigma\tau H\,. \qquad (*)$$

En effet, l'inclusion $\sigma\tau H \subset (\sigma H)(\tau H)$ est évidente (car $1 \in H$), et pour la réciproque, on écrit

$$\sigma h_1 \tau h_2 = \sigma\tau(\tau^{-1}h_1\tau)h_2 \in \sigma\tau H$$

puisque $\tau^{-1}h_1\tau \in H$. D'où $(\sigma H)(\tau H) \subset \sigma\tau H$.

Ceci nous donne une multiplication sur G/H, et (*) montre que $\overline{\sigma}\,\overline{\tau} = \overline{\sigma\tau}$. On en déduit facilement, comme dans le cas des modules, que l'on a bien une structure de groupe sur G/H. La démonstration que le noyau de p est H est également similaire à celle dans le cas des modules. □

Très souvent, nous utiliserons la lettre N pour un sous-groupe distingué de G (c'est parce qu'on dit parfois « groupe normal » au lieu de « groupe distingué »).

Proposition 140. *Soit N un sous-groupe distingué de G, et soit $\varphi\colon G \longrightarrow H$ un homomorphisme tel que $N \subset \ker\varphi$. Alors il existe un unique homomorphisme $\overline{\varphi}\colon G/N \longrightarrow H$ tel que $\overline{\varphi}(\overline{g}) = \varphi(g)$ pour tout $g \in G$. De plus, si φ est surjective et si $N = \ker(\varphi)$, alors $\overline{\varphi}$ est un isomorphisme, ce qu'on résume en*

$$G/\ker(\varphi) \cong \mathrm{Im}(\varphi)\,.$$

Démonstration. Comme pour les modules. □

Exemple 141. Vous connaissez sans doute l'homomorphisme dit « de signature »

$$\varepsilon\colon S_n \longrightarrow \{\pm 1\}\,,$$

et vous devez savoir que son noyau est appele le *groupe alterné* de degré n, noté A_n. On a donc

$$S_n/A_n \cong \{\pm 1\} \cong \mathbb{Z}/2\mathbb{Z}\,.$$

Le théorème de Lagrange

Comme promis, considérer les classes à gauche, c'est-à-dire les éléments de G/H, peut servir à d'autres choses qu'à former des quotients. Commençons par :

Lemme 142. *Soit xH et yH deux éléments de G/H. Si $xH \cap yH \neq \emptyset$, alors $xH = yH$.*

Démonstration. Supposons que $z \in xH \cap yH$. Alors $z = xh$ pour un $h_1 \in H$, et $z = yh_2$ pour un $h_2 \in H$, par définition. Donc $x = zh_1^{-1} = yh_2h_1^{-1} \in yH$; donc $xH \subset yH$, clairement ; et de manière symétrique, on a aussi $yH \subset xH$, d'où $xH = yH$. □

En corollaire, nous avons

Théorème 143 (Lagrange)**.** *Soit G un groupe fini, et soit H un sous-groupe de G. Alors l'ordre de H divise l'ordre de G. Plus précisément, nous avons*

$$\frac{|G|}{|H|} = |G/H|\,.$$

Démonstration. Si $g \in G$, on a bien sûr $g \in gH$, donc G est l'union de tous les ensembles gH, où g parcourt G. Mais d'après le lemme, cette union est *disjointe*. Si $k = |G/H|$, le groupe G est donc l'union disjointe de $x_1H, x_2H, \ldots, x_kH$, pour des $x_i \in G$ (non-uniques !), de sorte que

$$|G| = |x_1H| + \cdots + |x_kH| \,. \qquad (*)$$

Mais chaque ensemble gH est en bijection avec H : considérer l'application $H \longrightarrow gH$ définie par $h \mapsto gh$, d'inverse $x \mapsto g^{-1}x$. Donc $|gH| = |H|$ pour tout g, et en particulier $|x_iH| = |H|$ pour tout i. De l'égalité (*) on tire alors $|G| = k|H|$. □

Corollaire 144. *Si $g \in G$, alors l'ordre de g divise l'ordre de G. En particulier, on a $g^{|G|} = 1$.*

Démonstration. Le théorème de Lagrange appliqué à $H = \langle g \rangle$ donne la première phrase. Si k est l'ordre de g, on a donc $|G| = kd$, donc $g^{|G|} = (g^k)^d = 1^d = 1$. □

Corollaire 145 (Petit théorème de Fermat)**.** *Soit p un nombre premier, et $x \in \mathbb{Z}$. Alors $x^p \equiv x \bmod p$. Si p ne divise pas x, alors $x^{p-1} \equiv 1 \bmod p$.*

Démonstration. On prend $G = (\mathbb{Z}/p\mathbb{Z})^*$, qui est d'ordre $p - 1$. Si x n'est pas divisible par p, alors $\overline{x} \in G$, et donc $\overline{x}^{p-1} = \overline{1}$ par le corollaire précédent, ce qui s'écrit $x^{p-1} \equiv 1 \bmod p$. On obtient $x^p \equiv x \bmod p$ en multipliant par x.

Si par contre p divise x, alors $x \equiv 0 \bmod p$, donc certainement $x^p \equiv 0 \equiv x \bmod p$. □

Toujours en guise d'application du théorème de Lagrange, voyons maintenant les sous-groupes d'un groupe cyclique fini d'ordre n : par Lagrange, l'ordre d'un tel sous-groupe doit diviser n, et nous allons voir que réciproquement, pour tout diviseur de n, il y a un sous-groupe de cet ordre, et même un seul. Nous avons déjà vu une partie de l'énoncé dans un autre exercice sur les modules, mais ici on va être plus précis (et c'est à retenir).

Théorème 146. *Soit $G = \langle g \rangle$ un groupe cyclique d'ordre n (en notation multiplicative). Soit d un diviseur de n, et écrivons $n = de$. Alors G possède un unique sous-groupe H_d d'ordre d, et de plus*

$$H_d = \langle g^e \rangle = \{x \in G \mid x^d = 1\}.$$

Démonstration. Commençons par montrer que

$$\langle g^e \rangle = \{x \in G \mid x^d = 1\}.$$

En effet, un $x \in G$ peut s'écrire $x = g^k$, par définition. Alors $x^d = g^{kd}$, de sorte que $x^d = 1$ équivaut à $g^{kd} = 1$, c'est-à-dire à $n|kd$. Mais $de|dk$ si et seulement si $e|k$; et si $k = e\ell$, alors $g^k = (g^e)^\ell$, et $x \in \langle g^e \rangle$. Finalement, on constate que $x^d = 1$ équivaut à $x \in \langle g^e \rangle$, d'où l'égalité. On note H_d pour le groupe en question.

Le groupe H_d est cyclique, engendré par g^e, donc son ordre est le plus petit entier k tel que $(g^e)^k = 1$ (lemme du chapitre 1). Mais $g^{ek} = 1$ si et seulement si $n|ek$, donc $d|k$. L'entier k est un multiple de d, donc $k \geq d$; mais $(g^e)^d = g^n = 1$ par Lagrange, donc par minimalité on a $k = d$. On a montré que H_d est d'ordre d.

Montrons enfin l'unicité. Soit H un sous-groupe de G d'ordre d. Alors pour tout $x \in H$, on a $x^d = 1$ par Lagrange ; ceci montre $H \subset H_d$, et par égalité des ordres, $H = H_d$. □

§4.2 Actions

Définition 147. Soit X un ensemble. Alors $S(X)$ est le groupe de toutes les bijections $X \longrightarrow X$, avec la composition $\circ$ pour loi de groupe, et l'identité pour élément neutre. On l'appelle le *groupe symétrique de* X. Lorsque $X = \{1, 2, \ldots, n\}$, on écrit S_n pour $S(X)$.

Définition 148. On dit que le groupe G *agit* sur l'ensemble X lorsqu'il existe un homomorphisme $\varphi : G \longrightarrow S(X)$. On dira parfois que φ est une *action* de G sur X.

En général, on va écrire $g \cdot x$ au lieu de $\varphi(g)(x)$, pour $g \in G$, $x \in X$ (et φ est alors sous-entendu, mais c'est toujours comme ça). Avec cette notation, on a

$$g \cdot (h \cdot x) = (gh) \cdot x\,, \qquad 1 \cdot x = x\,.$$

Il est facile de montrer que, si on a une application $G \times X \longrightarrow X$ notée $(g, x) \mapsto g \cdot x$, et si les deux propriétés ci-dessus sont vérifiées, alors $g \cdot x = \varphi(g)(x)$ où φ est une action unique. Bref, on aurait pu donner une autre définition, et dans certains livres, c'est ce que vous trouverez.

Exemple 149. Le groupe diédral D_8 agit sur $\mathbb{R}^2$. Le groupe S_n agit sur $\{1, 2, \cdots, n\}$.

Exemple 150. On peut prendre $X = G$, avec l'action $g \cdot x = gx$ (multiplication dans G). On dira « l'action de G sur lui-même par multiplication à gauche ». Notons également que, si H est un sous-groupe de H, alors H agit sur G par multiplication, par restriction de l'action précédente.

Exemple 151. Prenons encore $X = G$, avec cette fois-ci $g \cdot x = gxg^{-1}$. Vérifions qu'il s'agit bien d'une action : on a

$$g \cdot (h \cdot x) = g \cdot (hxh^{-1}) = ghxh^{-1}g^{-1} = (gh)x(gh)^{-1} = (gh) \cdot x\,.$$

D'autres part $1 \cdot x = x$ est évident. C'est bien une action, et nous parlerons de « l'action de G sur lui-même par conjugaison ».

Exemple 152. Prenons $X = \mathscr{P}(G)$, l'ensemble des parties $S \subset G$. Alors on définit une action de G sur X en posant

$$g \cdot S = \{gs \mid s \in S\}\,.$$

Si H est un sous-groupe de G, alors $G/H \subset \mathscr{P}(G)$, et l'action de G sur $\mathscr{P}(G)$ envoie G/H dans lui-même ; en effet si $xH \in G/H$ alors

$$g \cdot (xH) = gxH \in G/H\,,$$

comme on le vérifie (il y a un tout petit calcul à faire ici). Il y a donc une action de G sur G/H, que appellerons *canonique*, et qui est très importante. Si on utilise la notation « avec les barres », alors on a

$$g \cdot \overline{x} = \overline{gx}\,.$$

Définition 153. Supposons que G agit sur X, et soit $x \in X$. Alors *l'orbite de* x, notée $\mathrm{orb}(x)$, est

$$\mathrm{orb}(x) = \{g \cdot x \mid g \in G\}\,.$$

Exemple 154. Dans l'action de G sur lui-même par conjugaison, les orbites s'appellent les *classes de conjugaison*. Elles jouent un grand rôle dans les deux chapitres suivants.

Lemme 155. *Si* $\mathrm{orb}(x) \cap \mathrm{orb}(y) \neq \emptyset$, *alors* $\mathrm{orb}(x) = \mathrm{orb}(y)$. *En conséquence,* X *est l'union disjointe des différentes orbites de* G.

Démonstration. Laissé en exercice, très similaire au lemme ci-dessus sur les classes à gauche. □

Remarque. En fait, la condition $x \in \mathrm{orb}(y)$ est symétrique en x et y, puisque $x = g \cdot y$ équivaut à $y = g^{-1} \cdot x$. On voit facilement qu'elle est aussi transitive, et bien sûr réflexive, en d'autres termes c'est une relation d'équivalence. Les orbites sont les classes d'équivalence, et voilà qui explique pourquoi elles sont disjointes.

Dans le cas où X est fini, on peut donc trouver des éléments $x_1, \ldots, x_k$ tels que

$$X = \mathrm{orb}(x_1) \cup \mathrm{orb}(x_2) \cup \cdots \cup \mathrm{orb}(x_k)\,,$$

une union disjointe. Les x_i ne sont pas uniques du tout : si on prend un $x \in \mathrm{orb}(x_i)$ quelconque, alors $\mathrm{orb}(x) = \mathrm{orb}(x_i)$ (puisque ces deux orbites contiennent toutes les deux l'élément x !).

Il est bien clair que l'action de G préserve les orbites, c'est-à-dire que si $x \in \mathrm{orb}(x_i)$, alors $g \cdot x \in \mathrm{orb}(x_i)$ pour tout $g \in G$. Donc une action de groupe peut souvent s'étudier « orbite par orbite ». On a d'ailleurs le vocabulaire suivant :

Définition 156. L'action de G sur X est dite *transitive* lorsqu'il n'y a qu'une seule orbite. Il est équivalent (mini-exo) de demander que pour tous $x, y \in X$, il existe un $g \in G$ tel que $g \cdot x = y$.

Ce qui précède montre que l'étude des actions de groupe peut en grande partie se ramener à l'étude des actions transitives.

Bref, nous avons saisi l'importance des orbites et de l'action induite sur elles. Or la structure des orbites est très bien comprise. Voyons ça.

Définition 157. Supposons que G agit sur X, et soit $x \in X$. Alors le *stabilisateur* de x pour cette action est

$$\mathrm{Stab}_G(x) = \{g \in G \mid g \cdot x = x\}\,.$$

C'est un sous-groupe de G.

Exemple 158. Reprenons l'exemple de l'action de G par conjugaison sur lui-même. Le stabilisateur de $x \in G$ est constitué des $g \in G$ tels que $gxg^{-1} = x$, ce qui revient à $gx = xg$. C'est donc le sous-groupe des éléments qui commutent avec x, et on l'appelle en général le *centralisateur de x dans G*.

Proposition 159. *On suppose que G agit sur X. Soit $x \in X$, et soit $H = \mathrm{Stab}_G(x)$. Alors l'application*

$$\theta\colon G/H \longrightarrow \mathrm{orb}(x)$$

$$\overline{g} = gH \mapsto g \cdot x$$

est bien définie, c'est une bijection, et de plus pour $\sigma \in G$ et $\overline{g} \in G/H$ on a

$$\theta(\sigma \cdot \overline{g}) = \sigma \cdot \theta(\overline{g})\,.$$

Ici le $\cdot$ fait référence, à gauche, à l'action de G sur G/H, et à droite, à l'action de G sur X. En d'autres termes, on a une identification de $\mathrm{orb}(x)$ avec G/H, qui est compatible avec les actions.

Démonstration de la proposition. Si $\overline{g_1} = \overline{g_2}$, c'est-à-dire $g_1H = g_2H$, alors il existe $h \in H$ tel que $g_2 = g_1h$. Donc

$$g_2 \cdot x = g_1 \cdot (h \cdot x) = g_1 \cdot x$$

puisque $h \in H = \mathrm{Stab}_G(x)$. Ainsi $g_1 \cdot x$ ne dépend que de $\overline{g_1}$, il est donc légitime de le noter $\theta(\overline{g_1})$, et l'application θ est bien définie.

Montrons que θ est surjective, donc prenons $y \in \mathrm{orb}(x)$. Alors $y = g \cdot x$ pour un $g \in G$ par définition, d'où $y = \theta(\overline{g})$.

Montrons que θ est injective, et supposons donc $\theta(\overline{g}_1) = \theta(\overline{g}_1)$, ce qui se traduit par

$$g_1 \cdot x = g_2 \cdot x\,,$$

d'où

$$g_2^{-1} g_1 \cdot x = x\,,$$

ce qui revient à $g_2^{-1} g_2 \in \mathrm{Stab}_G(x) = H$. On en déduit $g_1 H = g_2 H$, ou encore $\overline{g_1} = \overline{g_2}$. Ainsi, θ est une bijection.

Enfin, prenons $\sigma, g \in G$ et calculons

$$\theta(\sigma \cdot \overline{g}) = \theta(\overline{\sigma g}) = (\sigma g) \cdot x = \sigma \cdot (g \cdot x) = \sigma \cdot \theta(\overline{g})\,.$$

□

Cette proposition est riche de conséquences. Commençons par des résultats de « comptage » :

Corollaire 160. *On a*

$$|\,\mathrm{orb}(x)| = \frac{|G|}{|\,\mathrm{Stab}_G(x)|}\,.$$

En particulier, la taille d'une orbite est un diviseur de l'ordre de G.

Démonstration. C'est la proposition ci-dessus et le théorème de Lagrange. □

Et en conséquence :

Corollaire 161. *On suppose que le groupe fini G agit sur l'ensemble fini X, et on prend $x_1, x_2, \ldots, x_k$ tels que X soit l'union disjointe*

$$X = \mathrm{orb}(x_1) \cup \cdots \cup \mathrm{orb}(x_k)\,.$$

Alors

$$|X| = \sum_i \frac{|G|}{|\,\mathrm{Stab}_G(x_i)}\,.$$

On appelle parfois cette identité « l'équation aux classes ». Il paraît beaucoup moins important de la mémoriser que de mémoriser la proposition.

Démonstration. Bien sûr $|X|$ est la somme des nombres $|\,\mathrm{orb}(x_i)|$. Or, le corollaire précédent affirme que $|\,\mathrm{orb}(x_i)| = |G|/|\,\mathrm{Stab}_G(x_i)$. □

§4.3 Les théorèmes de Sylow

Définition 162. Soit G un groupe d'ordre n, et soit p un nombre premier. Écrivons $n = p^k m$ avec m premier à p. Alors un *p-Sylow* de G est un sous-groupe H d'ordre p^k.

(Si vous avez beaucoup de temps devant vous, vous pouvez dire « un sous-groupe de Sylow associé au nombre premier p » au lieu de dire « un p-Sylow ».)

Il n'est pas évident du tout, étant donné un groupe G arbitraire, qu'il possède un p-Sylow. Mais c'est ce qu'affirment, entre autres choses, les théorèmes de Sylow que nous allons démontrer.

Nous aurons besoin d'une notation : si H est un sous-groupe de G, et si $g \in G$, sans surprise on pose

$$Hg = \{hg \mid h \in H\}\,.$$

Le point crucial va être de considérer l'ensemble

$$X = \{A \subset G \text{ tel que } |A| = p^k\}\,.$$

Ici A n'est pas un sous-groupe, mais juste une partie du groupe G d'ordre $n = p^k m$, comme ci-dessus. On fait agir G sur X par

$$g \cdot A = \{ga \mid a \in A\}\,;$$

en d'autres termes, c'est l'action que l'on a déjà vue de G sur l'ensemble $\mathscr{P}(G)$ de ses parties, sauf qu'on se restreint aux parties de cardinal p^k.

On a alors le lemme suivant, qui détient la clef des théorèmes de Sylow.

Lemme 163. *Dans les notations ci-dessus, soit $A \in X$ et H son stabilisateur. Les conditions ci-dessous sont équivalentes.*

1. $|\mathrm{orb}(A)|$ *est premier à p.*
2. *il existe $g \in G$ tel que $A = Hg$.*
3. $\mathrm{orb}(A)$ *contient un unique p-Sylow, qui est un conjugué de H.*
4. *H est un p-Sylow de G.*

5. $|\operatorname{orb}(A)| = m$.

Démonstration. Supposons (1), et montrons (2). Pour cela, prenons $g \in A$ un élément quelconque. Pour $h \in H$, on a $hg \in hA = A$ puisque H est le stabilisateur de A. Donc $Hg \subset A$. Mais $|Hg| = |H|$, donc on en déduit que $|H| \leq |A| = p^k$. Mais en même temps, $|\operatorname{orb}(A)| = |G|/|H|$ est premier à p, ce qui veut dire que $|H| = p^k r$ avec $r|m$. Au total, on doit donc avoir $r = 1$ et $|H| = p^k$. Puisque $|Hg| = |H| = p^k$, on en déduit $Hg = A$.

Passons à (2) $\implies$ (3). Si $A = Hg$, posons $K = g^{-1}Hg = g^{-1}A$, de sorte que $|K| = |H| = |A| = p^k$, et H et K sont tous les deux des p-Sylow. Mais de plus, $K \in \operatorname{orb}(A)$ par construction, et en fait c'est le seul ensemble dans $\operatorname{orb}(A) = \operatorname{orb}(K)$ qui est un sous-groupe de G (si xK est un sous-groupe, il contient $1 \in G$ d'où $x \in K$ et $xK = K$). C'est donc, en particulier, le seul p-Sylow dans $\operatorname{orb}(A)$. Nous avons (3).

L'implication (3) $\implies$ (4) est triviale. Si on a (4), on écrit $|\operatorname{orb}(A)| = |G|/|H| = m$, donc on a (5). Et (5) $\implies$ (1) est encore trivial. □

Corollaire 164. *Il existe une bijection entre les p-Sylow de G et les orbites de cardinal premier à p dans X, donnée par $H \mapsto \operatorname{orb}(H)$.*

Démonstration. En effet, soit H un p-Sylow, et prenons $A = H$ dans la notation du lemme ; alors $\operatorname{Stab}_G(H) = H$, puisque la condition $gH = H$ équivaut à $g \in H$. Donc $\operatorname{orb}(H)$ est de cardinal premier à p, par l'implication (4) $\implies$ (1). Donc l'application $H \mapsto \operatorname{orb}(H)$ envoie bien le p-Sylow sur une orbite de cardinal premier à p.

Si $\operatorname{orb}(H) = \operatorname{orb}(K)$ où H et K sont tous les deux des p-Sylow, alors par la propriété (3) (la partie unicité, en fait), on a $H = K$; l'application $H \mapsto \operatorname{orb}(H)$ est donc injective.

Enfin, l'implication (1) $\implies$ (3) montre que toute orbite $\operatorname{orb}(A)$ de cardinal premier à p contient un p-Sylow K, mais alors $\operatorname{orb}(K) = \operatorname{orb}(A)$. Ainsi, l'application est également surjective, ce qui termine la démonstration. □

Il va donc être essentiel, pour montrer que G contient au moins un p-Sylow, de comprendre la taille des orbites dans X. Or, on a

$$|X| = \binom{p^k m}{p^k}$$

par définition. Et nous notons :

Lemme 165. *Si $pgcd(m,p) = 1$, où p est premier, alors pour tout $k \geq 0$ on a*

$$\binom{p^k m}{p^k} \equiv m \bmod p\,.$$

Démonstration. On peut le faire en raisonnant avec les nombres binomiaux, c'est un bon exercice. Mais on peut aussi le déduire du lemme précédent (ce n'est pas la fonction principale de ce lemme, mais ça marche).

En effet, prenons $G = \mathbb{Z}/n\mathbb{Z}$, avec $n = p^k m$. Alors G possède un unique p-Sylow, d'après le théorème 146. Donc le lemme ci-dessus avec son corollaire nous dit qu'il existe une unique orbite de G dans X dont le cardinal est premier à p, et que ce cardinal est m. Si les orbites sont $\mathrm{orb}(A_1), \ldots, \mathrm{orb}(A_k)$, avec $|\mathrm{orb}(A_1)| = m$ et $\mathrm{orb}(A_i)$ divisible par p pour $i \geq 2$, on a bien

$$|X| \equiv |\mathrm{orb}(A_1)| \bmod p\,.$$

□

Il est maintenant très facile d'établir :

Théorème 166 (Premier théorème de Sylow)**.** *Soit G un groupe fini et p un nombre premier. Alors G possède un p-Sylow.*

Démonstration. En vue du corollaire au premier lemme, il faut montrer qu'il existe une orbite de cardinal premier à p, dans l'action de G sur X. Supposons par l'absurde que, les orbites étant $\mathrm{orb}(A_1), \ldots, \mathrm{orb}(A_k)$, chaque nombre $|\mathrm{orb}(A_i)|$ soit divisible par p. Alors $|X| = \sum_i |\mathrm{orb}(A_i)|$ est également divisible par p. Or c'est impossible, puisque $|X|$ est congru à m modulo p, où $pgcd(m,p) = 1$. □

Théorème 167 (Deuxième théorème de Sylow)**.** *Soit H un p-Sylow de G, soit K un sous-groupe de G qui est un p-groupe, et soit H un p-Sylow de G. Alors il existe $g \in G$ tel que $gKg^{-1} \subset H$. Si K est lui-même un p-Sylow, on en déduit que $gKg^{-1} = H$, et donc que tous les p-Sylow sont conjugués entre eux.*

Rappelons qu'un groupe fini est appelé un *p-groupe* lorsque son ordre est une puissance de p.

Démonstration. Sous forme d'exercices. □

Théorème 168 (Troisième théorème de Sylow)**.** *Soit n_p le nombre de p-Sylow distincts du groupe G d'ordre $p^k m$, avec $pgcd(m,p) = 1$. Alors*

- *n_p divise m,*
- *$n_p \equiv 1 \ mod\ p$.*

Démonstration. Le premier point va être une conséquence du deuxième théorème de Sylow. En effet, soit S l'ensemble des p-Sylow de G, de sorte que S est de cardinal n_p par définition. Le groupe G agit sur S par conjugaison (il est clair que si H est un p-Sylow, alors gHg^{-1} aussi), et le théorème précédent nous dit que cette action est transitive (= ne possède qu'une seule orbite). Si H est un p-Sylow, notons N son stabilisateur dans cette action. On a alors

$$n_p = |S| = |\operatorname{orb}(H)| = |G|/|N| \,.$$

Une remarque simple est alors que $H \subset N$ (en d'autres termes, si $h \in H$ alors $hHh^{-1} = H$, ce qui est évident). Finalement on a $H \subset N \subset G$, avec H d'ordre p^k et G d'ordre $p^k m$. Par Lagrange, l'ordre de N est $p^k r$ avec $r|m$. On conclut effectivement que

$$n_p = \frac{m}{r} \,,$$

ou encore $m = r n_p$.

Passons au deuxième point. Par le corollaire au premier lemme, il y a n_p orbites de cardinal premier à p dans l'action de G sur X, et ces orbites sont toutes de cardinal m par le lemme lui-même. L'union (disjointe) de ces orbites contient donc mn_p éléments de X, alors que les autres orbites sont de cardinal divisible par p. On a donc $|X| \equiv mn_p \bmod p$. Mais on sait également que $|X| \equiv m \bmod p$, d'où

$$m \equiv mn_p \bmod p \,.$$

Puisque m est premier à p, l'élément $\overline{m}$ est inversible dans $\mathbb{Z}/p\mathbb{Z}$, donc on peut simplifier par m pour obtenir $\overline{n_p} = \overline{1}$ dans $\mathbb{Z}/p\mathbb{Z}$. □

Chapitre 5

Représentations & caractères

Dans ce chapitre, nous dirons simplement « espace vectoriel » pour « espace vectoriel de dimension finie sur $\mathbb{C}$ ». Lorsque V est un espace vectoriel, nous noterons $GL(V)$ le groupe des automorphismes de V. Dès lors que l'on a choisi une base de V, nous l'utiliserons pour identifier $GL(V)$ avec $GL_n(\mathbb{C})$ (où $n = \dim V$), le groupe des matrices inversibles de taille $n \times n$.

Dans tout le chapitre, G va désigner un groupe fini. Il y aura un seul exemple, bien isolé, dans lequel on considèrera un groupe G infini.

§5.1 Représentations

Soit G un groupe. Dans la quasi-totalité des cas, on supposera que G est fini.

Définition 169. Soit V un espace vectoriel. Une *représentation linéaire de G dans V* est un homomorphisme

$$\rho : G \longrightarrow GL(V)$$

$$x \mapsto \rho_x \, .$$

Un homomorphisme $G \longrightarrow GL_n(\mathbb{C})$ sera considéré comme une représentation de G dans $\mathbb{C}^n$, et nous parlerons de *représentation matricielle* de G.

Voici une première remarque qui ne sera pas surprenante. Si ρ est une représentation de G dans V, notons $x \cdot v$ au lieu de $\rho_x(v)$. Alors on a pour $x, y \in G$, $v \in V$

$$x \cdot (y \cdot v) = (xy) \cdot v\,, \quad 1 \cdot v = v\,,$$

c'est-à-dire que l'on a une action de G sur V, et de plus

$$x \cdot (\lambda v) = \lambda(x \cdot v)\,, \quad x \cdot (v + w) = x \cdot v + x \cdot w\,,$$

puisque ρ_x est linéaire. Réciproquement, on voit sans difficulté que, si l'on a une application $G \times V \longrightarrow V$ notée $(x, v) \mapsto x \cdot v$ avec toutes ces propriétés, alors en notant $\rho_x : v \mapsto x \cdot v$, on obtient bel et bien une représentation ρ. En bref, comme dans le cas des actions (sur des ensembles), on a le choix entre deux définitions.

Dans ce chapitre, nous utiliserons plus souvent la notation ρ que la notation $x \cdot v$, puisque nous allons être amenés à considérer *toutes* les représentations d'un groupe donné (et on ne peut alors pas se contenter d'une notation dans laquelle l'homomorphisme $G \longrightarrow GL(V)$ est anonyme).

Définition 170. Soit ρ une représentation de G dans V et ρ' une représentation de G dans W. Une application $\mathbb{C}$-linéaire $\varphi : V \longrightarrow W$ est dite *G-linéaire* lorsque, pour tout $x \in G$, on a $\varphi \circ \rho_x = \rho'_x \circ \varphi$, comme dans le diagramme commutatif suivant :

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & W \\ {\scriptstyle \rho_x}\downarrow & & \downarrow{\scriptstyle \rho'_x} \\ V & \xrightarrow{\varphi} & W \end{array}$$

Avec la notation alternative, on a $\varphi(x \cdot v) = x \cdot \varphi(v)$ pour $x \in G$, $v \in V$.

Lorsque φ est un isomorphisme entre les espaces vectoriels V et W, sa réciproque φ^{-1} est également G-linéaire, et on dit que φ est un isomorphisme entre ρ et ρ'.

Pour s'entraîner avec le vocabulaire, montrons :

Lemme 171. *Soit G un groupe.*

1. *Toute représentation de G est isomorphe à une représentation matricielle.*

2. *Deux représentation matricielles* $\rho : G \longrightarrow GL_n(\mathbb{C})$ *et* $\rho' : G \longrightarrow GL_m(\mathbb{C})$ *sont isomorphes* $\Longleftrightarrow n = m$ *et il existe* $P \in GL_n(\mathbb{C})$ *inversible telle que*

$$\rho'_x = P\rho_x P^{-1}$$

pour tout $x \in G$.

Les objets que l'on étudie dans ce chapitre, les représentations, sont donc *les homomorphismes de* G *dans* $GL_n(\mathbb{C})$ *à conjugaison près.*

Démonstration. (1) Soit $\rho : G \longrightarrow GL(V)$ une représentation. Prenons une base de V, ce qui donne un isomorphisme $\varphi : V \longrightarrow \mathbb{C}^n$ où $n = \dim V$. Si on pose $\rho'_x = \varphi \circ \rho_x \circ \varphi^{-1}$, pour $x \in G$, alors $\rho'_x \in GL(\mathbb{C}^n) = GL_n(\mathbb{C})$. On vérifie tout de suite que ρ' est un homomorphisme puisque

$$\rho'_{xy} = \varphi \circ \rho_x \circ \rho_y \circ \varphi^{-1} = \varphi \circ \rho_x \circ \varphi^{-1} \circ \varphi \circ \rho_y \circ \varphi^{-1} = \rho'_x \circ \rho'_y,$$

donc ρ' est une représentation matricielle de G, et par définition φ est un isomorphisme entre ρ et ρ'.

(2) On suppose que ρ et ρ' sont isomorphes ; soit $\varphi : \mathbb{C}^n \longrightarrow \mathbb{C}^m$ un isomorphisme. Il est évident que $n = m$, et si P est la matrice de φ dans la base canonique, alors on a $\rho'_x = P\rho_x P^{-1}$.

Réciproquement, si une telle matrice existe, on définit φ comme étant l'application linéaire dont P est la matrice, et il est clair que c'est un isomorphisme entre ρ et ρ'. □

Exemple 172. Qu'est-ce qu'une représentation du groupe $G = \mathbb{Z}$? Ceci va être le seul et unique exemple avec G infini. A isomorphisme près, une représentation est donnée par un homomorphisme $\rho : \mathbb{Z} \longrightarrow GL_n(\mathbb{C})$, pour un certain n. Si on note $M = \rho(1)$, alors on doit avoir $\rho(k) = M^k$ pour tout $k \in \mathbb{Z}$, donc toute l'information est contenue dans la matrice M.

D'après le lemme, la représentation est déterminée à isomorphisme près par M à conjugaison près. Si l'on souhaite décrire toutes les représentations de $\mathbb{Z}$ à isomorphisme près, la réponse est donc : il y en a, en dimension n, exactement une pour chaque classe de conjugaison de $GL_n(\mathbb{C})$, et donc une pour chaque forme de Jordan (!) avec les valeurs propres non-nulles.

Le problème consistant à trouver toutes les représentations d'un groupe donné G à isomorphisme près est donc une généralisation du problème de la réduction des endomorphismes (celui-ci étant l'étude des classes de conjugaisons de $GL_n(\mathbb{C})$, essentiellement).

Voyons un exemple très simple :

Exemple 173. Prenons $G = \mathbb{Z}/d\mathbb{Z}$. Soit $\rho : G \longrightarrow GL_n(\mathbb{C})$ une représentation, et soit $M = \rho(\overline{1})$ comme ci-dessus. Cette fois-ci on a $M^k = \rho(\overline{k})$ pour tout entier k, et en particulier $M^d = \rho(\overline{d}) = \rho(\overline{0}) = I$. La matrice M est donc annulée par le polynôme $X^d - 1$, qui est scindé à racines simples, donc M est diagonalisable. Ainsi il existe P telle que PMP^{-1} est diagonale, avec des racines d-ièmes de l'unité sur la diagonale.

L'exemple précédent est censé illustrer le fait que, étant donné un groupe G, l'étude de ses représentations est un problème d'algèbre linéaire avec des contraintes imposées par les relations dans le groupe. Ici le fait que $\overline{d} = \overline{0}$ a fortement réduit les options pour la forme de Jordan de M.

On vous laisse réfléchir au fait que, si on avait pris $G = \mathbb{Z}/d\mathbb{Z} \times \mathbb{Z}/e\mathbb{Z}$, alors en posant $M = \rho(1, 0)$ et $N = \rho(0, 1)$, on doit avoir $MN = NM$ (en plus de $M^d = I = N^e$). Pour comprendre ρ, il faut donc se demander, en gros, si on peut diagonaliser simultanément M et N... (On va voir que la réponse est oui, et plus généralement le cas où G est abélien est parfaitement sous contrôle.)

Avant de démontrer des théorèmes sur les représentations, nous allons commencer par montrer qu'il existe une algèbre sur $\mathbb{C}$, notée $\mathbb{C}[G]$, telle que les $\mathbb{C}[G]$-modules (au sens des chapitres précédents) sont exactement les représentations de G. Nous n'allons pas étudier $\mathbb{C}[G]$ très en détail ; le but est simplement de se convaincre que ce chapitre est bel et bien une continuation de notre étude des modules, et par ailleurs, nous allons pouvoir économiser les définitions, puisque vous savez ce qu'est un sous-module, un module produit, etc.

Fixons donc le groupe fini G. On définit d'abord $\mathbb{C}[G]$ comme étant l'espace vectoriel des fonctions $\alpha : G \longrightarrow \mathbb{C}$. (La définition de la multiplication va attendre un tout petit peu.) En particulier, pour $g \in G$ nous allons considérer

l'élément $g^* \in \mathbb{C}[G]$ (la notation n'est pas standard) défini par

$$g^*(x) = \begin{cases} 1 \text{ si } x = g\,, \\ 0 \text{ sinon}\,. \end{cases}$$

Mais alors, une remarque indispensable est que, pour tout $\alpha \in \mathbb{C}[G]$, on a

$$\alpha = \sum_{g \in G} \alpha(g) g^*\,.$$

En effet, pour vérifier cette égalité de fonctions, il suffit de prendre $x \in G$ et d'évaluer les deux côtés ; on obtient bien $\alpha(x)$ dans les deux cas.

Les éléments g^*, pour $g \in G$, forment donc une base de $\mathbb{C}[G]$ (il est évident que c'est une famille libre). Pour ce qui est de la multiplication, nous souhaitons avoir

$$(gh)^* = g^* h^*\,.$$

La seule façon de faire ça est de définir, pour $\alpha, \beta \in \mathbb{C}[G]$, le produit $\alpha\beta \in \mathbb{C}[G]$ par :

$$(\alpha\beta)(x) = \sum_{g \in G} \alpha(g)\beta(g^{-1}x)\,.$$

(On dit parfois que c'est le *produit de convolution de α et β*). Il est très simple de vérifier avec cette formule, pour $\alpha = g^*$ et $\beta = h^*$, que $g^* h^* = (gh)^*$. Ce qui est beaucoup moins drôle, et que nous laisserons de côté, c'est de montrer que cette multiplication sur $\mathbb{C}[G]$ est bien associative, et fait de $\mathbb{C}[G]$ une algèbre sur $\mathbb{C}$.

Voici un résumé, et il est édifiant de remarquer que c'est tout ce que nous utiliserons par la suite – la définition de $\mathbb{C}[G]$ ne servira plus jamais !

Lemme 174 (Résumé)**.** *L'algèbre $\mathbb{C}[G]$ possède une base formée des g^*, où $g \in G$, de sorte que la dimension de $\mathbb{C}[G]$ est $|G|$. La multiplication sur $\mathbb{C}[G]$ vérifie, pour $g, h \in G$, que $g^* h^* = (gh)^*$.*

Comme vous vous en doutez, on fait souvent l'abus d'écrire tout simplement g au lieu de g^*, de sorte que l'on voit G comme un sous-groupe de $\mathbb{C}[G]^\times$, et la multiplication sur $\mathbb{C}[G]$ est la seule qui « prolonge » celle de G. Mais dans ce cours, nous éviterons cet abus.

Exemple 175. Pour se convaincre que le lemme nous suffit pour travailler avec $\mathbb{C}[G]$, et que l'on n'a pas besoin de la définition, prenons $G = S_3$ et faisons des calculs simples. Prenons disons

$$\alpha = -2(1,3)^* + 7(1,2,3)^*, \qquad \beta = (1,3)^* + (1,2)^*.$$

Alors

$$\begin{aligned}\alpha\beta &= -2(1,3)^*(1,3)^* + 7(1,2,3)^*(1,3)^* - 2(1,3)^*(1,2)^* + 7(1,2,3)^*(1,2)^* \\ &= -2I^* + 7(2,3)^* - 2(1,2,3)^* + 7(1,3)^*.\end{aligned}$$

Ici I est l'élément neutre de S_3. Avec un peu d'habitude, on écrit simplement -2 au lieu de $-2I^*$.

Lemme 176. *Soit A une algèbre sur $\mathbb{C}$, et soit $\rho : G \longrightarrow A^\times$ un homomorphisme de groupes. Alors il existe un unique homomorphisme d'algèbres $\rho^* : \mathbb{C}[G] \longrightarrow A$ tel que $\rho^*(g^*) = \rho(g)$ pour $g \in G$.*

Démonstration. L'unicité est évidente, puisque les g^* forment une base de $\mathbb{C}[G]$: nous devons poser

$$\rho^*\left(\sum_{g\in G} \lambda_g g^*\right) = \sum_{g\in G} \lambda_g \rho(g).$$

Réciproquement, cette formule définit une application linéaire $\rho^* : \mathbb{C}[G] \longrightarrow A$ et nous devons vérifier qu'il s'agit d'un homomorphisme d'anneaux, c'est-à-dire que $\rho^*(\alpha\beta) = \rho^*(\alpha)\rho^*(\beta)$. Par bilinéarité, il suffit de vérifier ça pour $\alpha = g^*$ et $\beta = h^*$, et dans ce cas

$$\rho^*(g^*h^*) = \rho^*((gh)^*) = \rho(gh) = \rho(g)\rho(h) = \rho^*(g^*)\rho^*(h^*).$$

Enfin, puisque $\rho(1) = 1$ (l'élément neutre de A), il est clair que ρ^* est bien un homomorphisme d'algèbres. □

Corollaire 177. *Soit V un espace vectoriel. Alors une représentation ρ de G dans V définit une structure de $\mathbb{C}[G]$-module sur V. Toute structure de $\mathbb{C}[G]$-module sur V est obtenue de la sorte, pour un ρ unique. Une application linéaire $V \longrightarrow W$ est G-linéaire (au sens de la définition donnée dans ce chapitre) si et seulement si elle est $\mathbb{C}[G]$-linéaire (au sens général du chapitre 1).*

Démonstration. Une représentation $\rho : G \longrightarrow GL(V) = \mathrm{End}_{\mathbb{C}}(V)^\times$ se « prolonge », d'après le lemme, en un homomorphisme $\rho^* : \mathbb{C}[G] \longrightarrow \mathrm{End}_{\mathbb{C}}(V)$, ce qui donne bien une structure de $\mathbb{C}[G]$-module sur V. Réciproquement, tout homomorphisme tel que ρ^* donne naissance à l'homomorphisme de groupes

$$\rho : G \longrightarrow \mathbb{C}[G]^\times \longrightarrow \mathrm{End}_{\mathbb{C}}(V)^\times = GL(V)\,.$$

Le reste du corollaire est assez évident, et on vous le laisse en exercice. □

Finalement, nous voyons que dans ce chapitre, nous étudions les $\mathbb{C}[G]$-modules et rien d'autre. D'ailleurs on dira souvent « soit V un $\mathbb{C}[G]$-module » au lieu de « soit ρ une représentation de G dans V », lorsque l'énoncé qui suit s'y prête mieux. (Tous les $\mathbb{C}[G]$-modules du chapitre sont supposés de dimension finie sur $\mathbb{C}$.)

En tout cas, vous connaissez les notions de sous-module, sommes directes de modules, produits, quotients... qui s'appliquent donc aux représentations. Mais les traductions sont de toute façon faciles. Par exemple, supposons que ρ est une représentation de G dans V, et que $U \subset V$. Alors U est un sous-$\mathbb{C}[G]$-module si et seulement si c'est un sous-espace vectoriel tel que $\rho_x(U) \subset U$ pour tout $x \in G$. Si U_1 et U_2 sont deux sous-$\mathbb{C}[G]$-modules de V, vous connaissez la définition de la notation $V = U_1 \oplus U_2$; si $\rho_i : G \longrightarrow GL(U_i)$ est l'homomorphisme correspondant pour $i = 1, 2$, on n'hésitera pas à écrire $\rho = \rho_1 \oplus \rho_2$ dans ce cas.

Soyons plus concrets ici. Prenons des bases de U_1 et U_2, de sorte que l'on puisse voir $\rho_i : G \longrightarrow GL_{d_i}(\mathbb{C})$ comme une représentation matricielle (ici $d_i = \dim U_i$). Alors $\rho = \rho_1 \oplus \rho_2$ signifie qu'il existe une base de V dans laquelle on a, en écriture par blocs :

$$\rho(x) = \begin{pmatrix} \rho_1(x) & 0 \\ 0 & \rho_2(x) \end{pmatrix},$$

pour $x \in G$.

Exemple 178. Retournons à l'exemple 173, donc à $G = \mathbb{Z}/d\mathbb{Z}$. On va complètement décrire toutes les représentations de G à isomorphisme près, et en

un sens, tout le reste de ce chapitre va consister en une généralisation à un groupe G quelconque des observations que nous allons faire.

Posons $\omega = \exp(2i\pi/d)$, de sorte que les racines d-ièmes de l'unité sont $1, \omega, \omega^2,$ Pour chaque $0 \leq k < d$, définissons

$$\rho_k : G \longrightarrow \mathbb{C}^\times$$

par $\rho_k(\overline{n}) = \omega^{kn}$, ce qui est bien défini. Pour $k = 1$, l'homomorphisme ρ_1 est un isomorphisme entre $G = \mathbb{Z}/d\mathbb{Z}$ et le groupes des racines d-ièmes de l'unité, qui envoie $\overline{1}$ sur ω, et pour k quelconque et $x \in G$, on a $\rho_k(x) = \rho_1(x)^k$.

Chaque ρ_k définit donc un $\mathbb{C}[G]$-module de dimension 1 sur $\mathbb{C}$; pour des raisons de dimension, ce module est donc *simple*.

Prenons maintenant une représentation quelconque ρ de G dans V. Comme on l'a vu dans l'exemple 173, l'élément $\rho(1) \in GL(V)$ vérifie $\rho(1)^d = I$. On peut donc diagonaliser $\rho(1)$, et ses valeurs propres sont des racines d-ièmes de l'unité ; c'est-à-dire qu'on peut trouver une base de V dans laquelle la matrice de $\rho(1)$ est

$$M = \begin{pmatrix} \omega^{k_1} & & & \\ & \omega^{k_2} & & \\ & & \ddots & \\ & & & \omega^{k_s} \end{pmatrix}.$$

(Ici $s = \dim V$.) Mais alors (dans cette base) on a $\rho(\overline{n}) = \rho(1)^n = M^n =$

$$\begin{pmatrix} \omega^{k_1 n} & & & \\ & \omega^{k_2 n} & & \\ & & \ddots & \\ & & & \omega^{k_s n} \end{pmatrix} = \begin{pmatrix} \rho_{k_1}(\overline{n}) & & & \\ & \rho_{k_2}(\overline{n}) & & \\ & & \ddots & \\ & & & \rho_{k_s}(\overline{n}) \end{pmatrix}.$$

Conclusion : $\rho = \rho_{k_1} \oplus \rho_{k_2} \oplus \cdots \oplus \rho_{k_s}$. Donc *tout $\mathbb{C}[G]$-module est une somme directe de $\mathbb{C}[G]$-modules simples.* De plus, ces modules simples sont à choisir parmi les modules ρ_k pour $0 \leq k < d$, et en conséquence il n'y a pas d'autres modules simples. En particulier *il n'y a qu'un nombre fini de $\mathbb{C}[G]$-modules simples.*

Les deux phrases en italique restent vraies pour un groupe fini G quelconque, comme nous allons le voir.

Avant de conclure ces préliminaires, ajoutons que pour tout groupe G, nous avons toujours au moins les représentations suivantes :

Définition 179. L'homomorphisme trivial $G \longrightarrow \mathbb{C}^\times = GL_1(\mathbb{C})$, qui vérifie $x \mapsto 1$ pour tout $x \in G$, est appelé la *représentation triviale* de G.

Définition 180. La *représentation régulière* de G est la représentation ρ qui correspond au module « régulier » $\mathbb{C}[G]$[1], que l'on notera tout simplement $\mathbb{C}[G]$ en général. On rappelle qu'il s'agit de l'anneau $\mathbb{C}[G]$ vu comme module sur lui-même, c'est-à-dire avec $a \cdot v = av$ pour $a, v \in \mathbb{C}[G]$. Concrètement (et nous allons voir ça dans les exercices), il y a une base formée des g^* pour $g \in G$, et on a

$$\rho_x(g^*) = (x^*) \cdot (g^*) = (x^*)(g^*) = (xg)^* .$$

Avec l'abus de notation qui consiste à oublier les astérisques, on a donc $x \cdot g = xg$.

§5.2 Semi-simplicité

Définition 181. Soit V un espace vectoriel. Un *produit scalaire hermitien*, ou produit hermitien, est une fonction

$$h \colon V \times V \longrightarrow \mathbb{C} ,$$

avec les propriétés suivantes pour $v, w \in V$:

1. l'application $x \mapsto h(v, x)$ est linéaire,
2. $h(v, w) = \overline{h(w, v)}$ (le conjugué complexe), en particulier $h(v, v)$ est réel,
3. $h(v, v) \geq 0$,
4. $h(v, v) = 0 \iff v = 0$.

Sur un espace vectoriel V donné, il existe toujours au moins un produit hermitien. En effet, il suffit de prendre une base $e_1, \ldots, e_n$ et de poser

$$h\left(\sum_i v_i e_i, \sum_i w_i e_i\right) = \sum_i \overline{v}_i w_i .$$

Lemme 182. *Soit G un groupe fini, et soit ρ une représentation de G dans V. Alors il existe un produit scalaire hermitien h sur V qui est G-invariant, dans le sens où*

$$h(\rho_x(v), \rho_x(w)) = h(v, w)$$

pour tout $x \in G, v, w, \in V$.

Démonstration. On prend tout d'abord un produit hermitien quelconque h_0, et puis on définit h comme étant « la moyenne de tous les transformés de h_0 par les éléments de G », c'est-à-dire concrètement

$$h(v, w) = \frac{1}{|G|} \sum_{x \in G} h_0(\rho_x(v), \rho_x(w)) \,.$$

Une fois qu'on a eu cette bonne idée (qui date de la fin du 19e siècle), les vérifications sont évidentes, et on vous les laisse à titre d'exercice. □

Corollaire 183. *Soit V un $\mathbb{C}[G]$-module, et soit $U_1 \subset V$ un sous-module. Alors il existe un autre sous-module U_2 tel que $V = U_1 \oplus U_2$.*

Démonstration. Soit h un produit hermitien G-invariant comme dans le lemme. Posons alors

$$U_2 = U_1^{\perp} = \{v \in V \mid h(u, v) = 0 \text{ pour tout } u \in U_1\} \,.$$

Alors U_2 est un sous-espace vectoriel de V, et c'est même un sous-$\mathbb{C}[G]$-module : en effet si $v \in U_2$ alors on écrit pour $u \in U_1$ et $x \in G$ que

$$h(u, \rho_x(v)) = h(\rho_{x^{-1}}(u), v) = 0$$

puisque $\rho_{x^{-1}}(u) \in U_1$. Donc $\rho_x(v) \in U_2$.

Montrons maintenant que $U_1 \cap U_2 = \{0\}$, ce qui est tout simple : en effet si v appartient à cette intersection, on a $h(v, v) = 0$, d'où $v = 0$.

Et enfin, montrons que $\dim U_1 + \dim U_2 = \dim V$ (c'est un classique de la théorie des produits hermitiens, évidemment, mais l'argument qui suit n'est pas le plus classique). On prend une base $e_1, \ldots, e_d$ de U_1, que l'on complète

en une base $e_1, \ldots, e_n$ de V. On veut montrer que la dimension de U_2 est $n-d$. Considérons, pour ceci, l'application linéaire $\varphi : V \longrightarrow \mathbb{C}^n$ définie par

$$\varphi(v) = (h(e_1, v), h(e_2, v), \ldots, h(e_n, v)) .$$

Le noyau de φ est constitué des v tels que $h(e_i, v) = 0$ pour tout i, ce qui revient clairement à $h(w, v) = 0$ pour tout w ; pour $w = v$ on voit que cette condition implique $v = 0$. Donc φ est injective, donc c'est un isomorphisme.

L'espace U_2, qui est constitué des v tels que $h(e_i, v) = 0$ pour $1 \leq i \leq d$, est alors par définition égal à $\varphi^{-1}(\{0\} \times \mathbb{C}^{n-d})$. Il est donc bien de dimension $n - d$. □

Corollaire 184. *Soit V un $\mathbb{C}[G]$-module. Alors V est simple $\Longleftrightarrow$ V est indécomposable.*

Démonstration. On a toujours simple $\Longrightarrow$ indécomposable. Réciproquement, si V est indécomposable, soit $U \subset V$ un sous-module. Par le corollaire précédent, il existe un sous-module U' tel que $V = U \oplus U'$. Mais puisque V est indécomposable, on doit avoir ou bien $U = \{0\}$, ou bien $U' = \{0\}$ et dans ce cas $U = V$. □

Définition 185. On dit qu'un $\mathbb{C}[G]$-module est *irréductible* lorsqu'il est simple (ou indécomposable, ce qui revient au même). On parle aussi de *représentation irréductible* de G.

Théorème 186. *Soit V un $\mathbb{C}[G]$-module non-nul. Alors V se décompose en somme directe de modules irréductibles.*

On dit que V est *semi-simple*, d'où le titre de cette partie.

Démonstration. C'est très simple par récurrence sur $\dim V$. Si $\dim V = 1$, alors V est certainement irréductible. Supposons alors le théorème vrai pour tous les modules de dimension $< \dim V$. Si V est lui-même irréductible, il n'y a rien à montrer. Sinon, il possède un sous-module U avec $U \neq \{0\}$ et $U \neq V$. D'après le corollaire 183, on a $V = U \oplus U'$ pour un certain sous-module U', et maintenant on a $\dim U < \dim V$ et $\dim U' < \dim V$. On peut donc décomposer U et U' en somme directe d'irréductibles, donc V aussi. □

§5.3 Le lemme de Schur

Ici G est un groupe.

Définition 187. Soient V et W des $\mathbb{C}[G]$-modules. On note $\mathrm{Hom}_G(V,W)$ l'ensemble des applications $\mathbb{C}[G]$-linéaires $V \longrightarrow W$. C'est un espace vectoriel sur $\mathbb{C}$.

Lemme 188 (Schur)**.** *Soient V et W des $\mathbb{C}[G]$-modules irréductibles.*

1. *Toute application $\mathbb{C}[G]$-linéaire $V \longrightarrow W$ non-nulle est un isomorphisme. En particulier, si V et W ne sont pas isomorphes, on a $\mathrm{Hom}_G(V,W) = \{0\}$.*
2. *$\mathrm{Hom}_G(V,V) = \{\lambda I \mid \lambda \in \mathbb{C}\}$, où I désigne l'identité.*

Démonstration. Soit $f : V \longrightarrow W$ une application $\mathbb{C}[G]$-linéaire, avec f non-nulle. Alors son image $\mathrm{Im}(f)$ est un sous-module de W non-nul, et puisque W est irréductible, on en déduit $\mathrm{Im}(f) = W$, c'est-à-dire que f est surjective. De même, $\ker(f)$ est un sous-module de V différent de V, et puisque V est irréductible, on en déduit $\ker(f) = \{0\}$, c'est-à-dire que f est injective. Donc f est un isomorphisme entre V et W. Voilà qui montre le point (1).

Dans le cas où $V = W$, on prend une valeur propre λ de l'endomorphisme f : il en possède au moins une, puisqu'on est sur $\mathbb{C}$. Alors $\ker(f - \lambda I)$ est un sous-module de V non-nul, donc $\ker(f - \lambda I) = V$, toujours parce que V est irréductible. D'où $f = \lambda I$. □

Voici une notation pour énoncer une conséquence très utile :

Définition 189. Soient V et W des $\mathbb{C}[G]$-modules. On va noter

$$(V,W) = \dim \mathrm{Hom}_G(V,W) \in \mathbb{N}\,.$$

Parfois on notera $(V,W)_G$ s'il y a une ambiguïté sur le groupe G considéré.

Corollaire 190. *Avec V et W irréductibles, on a*

$$(V,W) = \begin{cases} 1 \text{ si } V \cong W\,, \\ 0 \text{ sinon}\,. \end{cases}$$

Démonstration. On a certainement $\dim \mathrm{Hom}_G(V, W) = 0$ si V et W ne sont pas isomorphes, par le premier point du lemme. Par contre, s'il existe un isomorphisme $V \cong W$ de $\mathbb{C}[G]$-modules, alors on en déduit un isomorphisme $\mathrm{Hom}_G(V, W)$ $\mathrm{Hom}_G(V, V)$ d'espaces vectoriels. Or $\dim \mathrm{Hom}_G(V, V) = 1$ d'après le lemme. □

Lemme 191. *La notation ci-dessus a les propriétés suivantes :*

1. *Si $V \cong V'$ et $W \cong W'$, alors $(V, W) = (V', W')$,*
2. *$(V, W \oplus W') = (V, W) + (V, W')$,*
3. *$(V \oplus V', W) = (V, W) + (V', W)$,*
4. *$(V, W) = (W, V)$.*

Démonstration. Les trois premières sont presque évidentes, et on vous les laisse en exercice. Pour la (4), on note que la propriété est vraie si V et W sont irréductibles, par le dernier corollaire ; mais comme tout module est une somme directe de modules irréductibles, les points (2) et (3) permettent de montrer le (4) dans le cas général. □

On peut maintenant montrer :

Proposition 192. *La décomposition d'un $\mathbb{C}[G]$-module en somme d'irréductibles est unique. Plus précisément, supposons que*

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k\,,$$

et que

$$V = V_1' \oplus V_2' \oplus \cdots \oplus V_\ell'\,,$$

où les V_i et les V_i' sont irréductibles. Alors $k = \ell$, et après renumérotation si nécessaire, on a $V_i \cong V_i'$ pour $1 \leq i \leq k$.

Attention : la conclusion n'est pas une égalité $V_i = V_i'$, mais seulement un isomorphisme.

Démonstration. Soit S un $\mathbb{C}[G]$-module irréductible quelconque (par exemple S peut être l'un des V_i, ou bien l'un des V_i'). Alors

$$(V,S) = \left(\bigoplus_i V_i, S\right) = \sum_i (V_i, S) = \text{ le nombre d'indices } i \text{ tels que } V_i \cong S\,.$$

Mais on a également

$$(V,S) = \left(\bigoplus_i V_i', S\right) = \sum_i (V_i', S) = \text{ le nombre d'indices } i \text{ tels que } V_i' \cong S\,.$$

Il y a donc autant de modules isomorphes à S parmi les V_i que parmi les V_i'. Le résultat en découle. □

Corollaire 193. *Soit V un $\mathbb{C}[G]$-module, et soit U un sous-module de V qui est irréductible. Alors il n'y a qu'un nombre fini de possibilités pour U, à isomorphisme près. Plus précisément, si $V = V_1 \oplus \cdots \oplus V_k$ est une décomposition de V en somme de sous-modules irréductibles, alors U est isomorphe à l'un des V_i.*

Démonstration. On peut trouver des sous-modules irréductibles $U_2, U_3, \ldots U_\ell$ de V tels que

$$V = U \oplus U_2 \oplus U_3 \oplus \cdots \oplus U_\ell\,.$$

Il suffit alors d'appliquer la proposition (et au passage on voit $\ell = k$). □

Corollaire 194. *Soit G un groupe fini et S un $\mathbb{C}[G]$-module irréductible. Alors il n'y a qu'un nombre fini de possibilités pour S, à isomorphisme près.*

Démonstration. Nous allons voir dans les exercices, et nous reverrons d'une autre manière ci-dessous, que S est isomorphe à un sous-module de la représentation régulière de G. □

§5.4 Calculs explicites

On fixe un groupe fini G. Notre but est ici de rendre le calcul du nombre (V, W) beaucoup plus explicite.

Lemme 195. *Soient V et W des $\mathbb{C}[G]$-modules. Alors l'espace vectoriel $\mathrm{Hom}(V,W)$ des applications linéaires de V dans W possède une structure de $\mathbb{C}[G]$-module. De plus, on peut identifier $\mathrm{Hom}_G(V,W)$ avec le sous-espace des points fixes, c'est-à-dire*

$$\mathrm{Hom}_G(V,W) = \{\varphi \in \mathrm{Hom}(V,W) \mid x \cdot \varphi = \varphi \text{ pour } x \in G\}\,.$$

Démonstration. On utilise la notation « avec un point », et on doit donc définir pour chaque $\varphi \in \mathrm{Hom}(V,W)$ un application $x \cdot \varphi \colon V \longrightarrow W$. On prend :

$$(x \cdot \varphi)(v) = x \cdot \varphi(x^{-1} \cdot v)\,.$$

Il faut vérifier que c'est bien une représentation de G, donc que $x \cdot (y \cdot \varphi) = (xy) \cdot \varphi$, et ainsi pour chaque $v \in V$ on calcule :

$$\big[x\cdot(y\cdot\varphi)\big](v) = x\cdot\big[(y\cdot\varphi)(x^{-1}\cdot v)\big] = xy\cdot\varphi(y^{-1}x^{-1}\cdot v) = xy\cdot\varphi((xy)^{-1}\cdot v) = \big[(xy)\cdot\varphi\big](v$$

Ça marche. On a évidemment $1\cdot\varphi = \varphi$, donc il y a bien une représentation de G dans $\mathrm{Hom}(V,W)$, ou ce qui revient au même, une structure de $\mathbb{C}[G]$-module.

La condition $x\cdot\varphi = \varphi$ revient à dire que pour tout $v \in V$ on a $x\cdot\varphi(x^{-1}\cdot v) = \varphi(v)$ ou encore $\varphi(x^{-1}\cdot v) = x^{-1}\cdot\varphi(v)$. Ceci est vrai pour tout $x \in G$ exactement lorsque φ est G-linéaire, c'est-à-dire lorsque $\varphi \in \mathrm{Hom}_G(V,W)$. □

Puisque nous nous intéressons à la dimension de $\mathrm{Hom}_G(V,W)$, le lemme nous apprend que nous sommes en train de faire un calcul de points fixes. Or on a aussi :

Lemme 196. *Soit ρ une représentation de G dans l'espace vectoriel M, et notons*

$$M^G = \{v \in M \mid \rho_x(v) = v \text{ pour tout } x \in G\}\,.$$

Soit $\pi \colon M \longrightarrow M$ définie par

$$\pi(v) = \frac{1}{|G|}\sum_{x\in G} \rho_x(v)\,.$$

Alors $\pi \circ \pi = \pi$, c'est-à-dire que π est un projecteur, et $\mathrm{Im}(\pi) = M^G$.

Démonstration. Si $v \in M^G$, il est clair que $\pi(v) = v$, donc $M^G \subset \mathrm{Im}(\pi)$. Réciproquement, on note que pour $y \in G$:

$$\rho_y(\pi(v)) = \frac{1}{|G|} \sum_{x \in G} \rho_{yx}(v) = \frac{1}{|G|} \sum_{x \in G} \rho_x(v) = \pi(v)\,,$$

donc $\pi(v) \in M^G$ et $\mathrm{Im}(\pi) \subset M^G$. Donc $\mathrm{Im}(\pi) = M^G$. Et puisqu'on vient de voir que π est l'identité sur M^G, on a $\pi(\pi(v)) = \pi(v)$ pour tout $v \in M$, soit $\pi \circ \pi = \pi$. □

Par ailleurs, il est très facile de voir que π est $\mathbb{C}[G]$-linéaire, mais on ne va même pas s'en servir. Ce qui nous intéresse c'est :

Corollaire 197. *La dimension de M^G est donnée par*

$$\dim M^G = \frac{1}{|G|} \sum_{x \in G} \mathrm{Trace}(\rho_x)\,.$$

Démonstration. D'après le lemme, la dimension en question est le rang de l'endomorphisme π. Or le rang d'un projecteur est égal à sa trace. Cette propriété peut-être méconnue des projecteurs se montre en notant que $M = \mathrm{Im}(\pi) \oplus \ker(\pi)$, et que si on prend une base de $\mathrm{Im}(\pi)$ et une base de $\ker(\pi)$, alors leur réunion est une base de M dans laquelle π a pour matrice

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Le nombre de 1 est à la fois le rang de π, et sa trace. Étant donnée la formule pour π, sa trace est bien obtenue par la somme proposée dans le corollaire. □

L'ingrédient qui nous manque pour terminer notre calcul explicite des nombres (V, W) est :

Lemme 198. *Soient ρ et ρ' des représentations de G dans V et W respectivement, et soit ρ'' la représentation de G dans* $\mathrm{Hom}(V,W)$ *qui en résulte. Alors pour tout $x \in G$ on a*

$$\mathrm{Trace}(\rho''_x) = \mathrm{Trace}(\rho_{x^{-1}})\,\mathrm{Trace}(\rho'_x)\,.$$

Démonstration. Soit $N = |G|$. On a $x^N = 1$, donc $\rho_x^N = 1$, et l'endomorphisme ρ_x est diagonalisable, de même que $\rho_{x^{-1}}$. Pour les mêmes raisons, on peut diagonaliser ρ'_x. Soit $e_1, \ldots, e_m$ une base de V dans laquelle $\rho_{x^{-1}}$ est diagonale avec $\mu_1, \ldots, \mu_m$ sur la diagonale, et soit $\varepsilon_1, \ldots, \varepsilon_n$ une base de W dans laquelle ρ'_x est diagonale avec $\lambda_1, \ldots, \lambda_n$ sur la diagonale.

Ayant choisi ces bases, on peut identifier $\mathrm{Hom}(V,W)$ avec l'espace vectoriel des matrices $n \times m$ à coefficients dans $\mathbb{C}$ (et on identifie également $\rho_{x^{-1}}$ et ρ'_x avec des matrices). Soit $m^{(ij)}$ la matrice $n \times m$ dont tous les coefficients sont nuls sauf celui sur la ligne i, dans la colonne j, qui vaut 1. Alors les $m^{(ij)}$ forment une base de $\mathrm{Hom}(V,W)$.

Utilisons la notation A_{ij} pour le coefficient de la matrice A sur la ligne i, dans la colonne j (par exemple $m^{(ij)}_{ij} = 1$). Alors la trace de l'endomorphisme ρ''_x sur l'espace $\mathrm{Hom}(V,W)$ est

$$\mathrm{Trace}(\rho''_x) = \sum_{i,j} \rho''_x(m^{(ij)})_{ij}\,.$$

Par définition on a $\rho''_x(m^{(ij)}) = \rho'_x \circ m^{(ij)} \circ \rho_{x^{-1}}$. Mais on vérifie que $m^{(ij)} \circ \rho_{x^{-1}} = \mu_j m^{(ij)}$, et $\rho'_x \circ m^{(ij)} = \lambda_i m^{(ij)}$ (les matrices ρ_x^{-1} et ρ'_x étant diagonales, c'est très simple). Finalement

$$\rho''_x(m^{(ij)})_{ij} = \lambda_i \mu_j\,.$$

Or λ_i est le i-ème coefficient sur la diagonale de ρ'_x, et μ_j est le j-ème coefficient sur la diagonale de $\rho_{x^{-1}}$. En symboles

$$\rho''_x(m^{(ij)})_{ij} = (\rho'_x)_{ii}(\rho_{x^{-1}})_{jj}\,.$$

En faisant la somme sur tous les i et tous les j, on obtient la formule annoncée. □

Théorème 199 (formule fondamentale)**.** *Soient ρ et ρ' des représentations de G dans V et W respectivement. Alors*

$$(V, W) = \frac{1}{|G|} \sum_{x \in G} \mathrm{Trace}(\rho_{x^{-1}}) \, \mathrm{Trace}(\rho'_x) \, .$$

Démonstration. Il s'agit de récapituler. On rappelle que $(V, W) = \dim \mathrm{Hom}_G(V, W)$. D'après le lemme 195, nous cherchons la dimension du sous-espace des points fixes dans la représentation $\mathrm{Hom}(V, W)$. Le corollaire 197 donne une formule pour cette dimension, en fonction de la trace des opérateurs ρ''_x, dans la notation du lemme 198. Et c'est ce dernier lemme qui exprime cette trace comme un produit. □

On peut écrire cette « formule fondamentale », à l'aide du lemme suivant :

Lemme 200. *Soit ρ une représentation de G. Alors pour $x \in G$ on a*

$$\mathrm{Trace}(\rho_{x^{-1}}) = \overline{\mathrm{Trace}(\rho_x)} \, .$$

Ici la barre désigne le conjugué complexe.

Démonstration. Soit $n = |G|$. L'élément x vérifie $x^n = 1$, donc $\rho_x^n = 1$, et l'endomorphisme ρ_x est diagonalisable, avec des racines n-ièmes de l'unité sur la diagonale, disons $\omega_1, \omega_2, \ldots$ Par suite, $\rho_{x^{-1}} = \rho_x^{-1}$ est diagonalisable, avec $\omega_1^{-1}, \omega_2^{-1}, \ldots$, sur la diagonale. Mais bien sûr, lorsque z est un nombre complexe de module 1, on a $z^{-1} = \overline{z}$. D'où

$$\mathrm{Trace}(\rho_{x^{-1}}) = \sum_i \omega_i^{-1} = \sum_i \overline{\omega_i} = \overline{\mathrm{Trace}(\rho_x)} \, .$$ □

Avec les hypothèses du théorème, on peut donc écrire

$$(V, W) = \frac{1}{|G|} \sum_{x \in G} \overline{\mathrm{Trace}(\rho_x)} \, \mathrm{Trace}(\rho'_x) \, .$$

§5.5 Caractères

G désigne un groupe fini.

Définition 201. Une *fonction centrale* sur G est une fonction $f : G \longrightarrow \mathbb{C}$ telle que $f(yxy^{-1}) = f(x)$ pour tous les $x, y \in G$. On note $\mathscr{C}(G)$ l'espace vectoriel de toutes fonctions centrales (il est visiblement de dimension finie).

Définition 202. Soit ρ une représentation de G dans V. Le *caractère* de ρ, ou de V, est la fonction centrale notée χ_ρ ou χ_V donnée par

$$\chi_\rho(x) = \chi_V(x) = \mathrm{Trace}(\rho_x)$$

pour tout $x \in G$. C'est bien une fonction centrale, puisque

$$\chi_\rho(yxy^{-1}) = \mathrm{Trace}(\rho_y \circ \rho_x \circ \rho_y^{-1}) = \mathrm{Trace}(\rho_x) = \chi_\rho(x)\,,$$

à cause de la formule bien connue $\mathrm{Trace}(ABA^{-1}) = \mathrm{Trace}(B)$.

Un *caractère irréductible* de G est une fonction centrale de la forme χ_V avec V irréductible.

Remarques.

1. Soit ρ resp. ρ' une représentation de G dans V resp. V'. Si ρ et ρ' sont isomorphes, alors $\chi_\rho = \chi_{\rho'}$. En effet, soit $\varphi : V \longrightarrow V'$ un isomorphisme, alors

   $$\chi_{\rho'}(x) = \mathrm{Trace}(\varphi \circ \rho_x \circ \varphi^{-1}) = \mathrm{Trace}(\rho_x) = \chi_\rho(x)\,.$$

   Ce qui est beaucoup plus surprenant, c'est que la réciproque est vraie : si $\chi_\rho = \chi_{\rho'}$ alors ρ et ρ' sont isomorphes, voir ci-dessous.
2. Il est clair que $\chi_{V\oplus W} = \chi_V + \chi_W$ (mini-exo), et par suite, tout caractère est une somme de caractères irréductibles.
3. On a $\chi_V(1) = \dim V$ (ici 1 est l'élément neutre de G). Dans le langage des caractères, on dit que $\chi_V(1)$ est le *degré* du caractère χ_V.

Exemple 203. Un exemple très basique est celui d'une représentation de dimension 1. En effet, en prenant une représentation matricielle $\rho : G \longrightarrow GL_1(\mathbb{C}) = \mathbb{C}^\times$, on a $\chi_\rho(x) = \rho(x)$, puisque la trace d'un élément de $GL_1(\mathbb{C})$ est l'élément lui-même. Historiquement, au 19e siècle, un « caractère de G » désignait un homomorphisme $G \longrightarrow \mathbb{C}^\times$, et plus tard, la définition a été étendue comme ci-dessus, en utilisant la trace des représentations. Dans la terminologie présente, un homomorphisme de G vers $\mathbb{C}^\times$ est exactement la même chose qu'un

caractère de degré 1. À partir de maintenant, on utilisera typiquement une notation comme $\chi\colon G \longrightarrow \mathbb{C}^\times$ pour un tel homomorphisme, et on ne fait pas de distinction entre une représentation de dimension 1 et son caractère.

Définition 204. Pour $f, g \in \mathscr{C}(G)$, on définit

$$(f, g) = \frac{1}{|G|} \sum_{x \in G} \overline{f(x)} g(x) \,.$$

Il est clair que ceci définit un produit scalaire hermitien sur $\mathscr{C}(G)$.

La « formule fondamentale » que nous avons obtenue peut alors se reformuler comme ceci :

Lemme 205. *Soient V et W des $\mathbb{C}[G]$-modules. Alors*

$$(V, W) = (\chi_V, \chi_W) \,. \qquad \square$$

La conséquence suivante est frappante :

Corollaire 206. *Soient V et W des $\mathbb{C}[G]$-modules. Alors V et W sont isomorphes si et seulement si $\chi_V = \chi_W$.*

Ainsi, un $\mathbb{C}[G]$-module est déterminé, à isomorphisme près, par son caractère ! Alors qu'on pourrait penser que, en ne regardant que la trace des matrices au lieu des matrices elles-mêmes, beaucoup d'information se perdait. Il n'en est rien. (Même s'il ne faut pas exagérer non plus : le corollaire ne dit pas que l'on peut reconstruire explicitement V à partir de χ_V.)

Démonstration. On a déjà fait remarquer que si V et W sont isomorphes, on a $\chi_V = \chi_W$, le plus étonnant est la réciproque : supposons que $\chi_V = \chi_W$. Il suffit, pour montrer que V et W sont isomorphes, de montrer que pour tout module irréductible S, on a $(V, S) = (W, S)$ (pourquoi ?). Mais $(V, S) = (\chi_V, \chi_S) = (\chi_W, \chi_S) = (W, S)$. $\square$

Un aspect agréable de ce résultat est que, du côté des caractères, on n'a pas des énoncés « à isomorphisme près », mais des résultats « tout court ».

Par exemple, nous savons qu'il n'y a qu'un nombre fini de modules irréductibles *à isomorphisme près* (corollaire 194) ; donc il n'y a qu'un nombre fini de caractères irréductibles (et puis c'est tout).

Terminons cette partie avec ce qui est essentiellement un résumé – mais c'est souvent sous cette forme que vous verrez les choses dans un livre sur le sujet, alors il est utile de donner cette formulation.

Théorème 207. *Soient χ_1 et χ_2 deux caractères irréductibles du groupe fini G. Alors nous avons la « première relation d'orthogonalité » :*

$$\frac{1}{|G|}\sum_{x\in G}\overline{\chi_1(x)}\chi_2(x) = \begin{cases} 1 \text{ si } \chi_1 = \chi_2\,, \\ 0 \text{ sinon}\,. \end{cases}$$

Démonstration. La somme à gauche est exactement (χ_1, χ_2), par définition. Soient V et W des représentations telles que $\chi_1 = \chi_V$ et $\chi_2 = \chi_W$. Alors $(\chi_1, \chi_2) = (V, W)$, et par le lemme de Schur ce nombre vaut 1 dans le cas où $V \cong W$, ce qui se produit exactement lorsque $\chi_1 = \chi_2$, et 0 sinon. □

§5.6 La décomposition de la représentation régulière

Voici un exemple d'utilisation des caractères. On va considérer la représentation régulière, c'est-à-dire le module $\mathbb{C}[G]$, et appeler χ_{reg} son caractère.

Lemme 208. *Pour $x \in G$, nous avons*

$$\chi_{reg}(x) = \begin{cases} 0 \text{ si } x \neq 1\,, \\ |G| \text{ si } x = 1\,. \end{cases}$$

Par suite, pour toute $f \in \mathscr{C}(G)$, on a

$$(\chi_{reg}, f) = f(1)\,.$$

Démonstration. Soit ρ la représentation de G qui correspond au module régulier $\mathbb{C}[G]$. Par définition $\chi_{reg}(x)$ est la trace de l'endomorphisme ρ_x. Rappelons que $\mathbb{C}[G]$ possède une base formée des g^* pour $g \in G$, avec la formule $\rho_x(g^*) = (xg)^*$. Mais alors, si $x \neq 1$, la matrice de ρ_x dans cette base

ne possède que des 0 sur la diagonale (en fait il y a un seul coefficient non-nul dans chaque colonne, qui est un 1, et qui n'est pas sur la diagonale) ; donc $\mathrm{Trace}(\rho_x) = 0$.

Pour $x = 1$, la matrice ρ_1 est l'identité, et sa trace est donc la dimension de $\mathbb{C}[G]$. La formule pour χ_{reg} est donc établie.

Si maintenant $f \in \mathscr{C}(G)$, on calcule

$$(\chi_{reg}, f) = \frac{1}{|G|} \sum_{x \in G} \overline{\chi_{reg}(x)} f(x) = \frac{\chi_{reg}(1) f(1)}{|G|} = f(1) \,.$$

□

Cet énoncé purement en termes de caractères a pour conséquence :

Théorème 209. *Soient $\chi_1, \cdots, \chi_s$ les caractères irréductibles de G. Alors*

$$\chi_{reg} = \sum_i \chi_i(1)\, \chi_i \,.$$

Si V_i est une représentation de G de caractère χ_i, on a

$$\mathbb{C}[G] \cong \bigoplus_i V_i^{\dim V_i} \,.$$

En d'autres termes, chaque représentation irréductible de G est contenue dans $\mathbb{C}[G]$ autant de fois que sa dimension.

Démonstration. Puisque χ_{reg} est un caractère, il s'écrit certainement

$$\chi_{reg} = \sum_i m_i \chi_i$$

pour des entiers m_i, et alors $m_i = (\chi_{reg}, \chi_i)$ par le lemme de Schur. Le lemme donne donc $m_i = \chi_i(1)$.

Pour montrer le deuxième énoncé, il suffit de montrer que les deux modules de part et d'autre du signe $\cong$ ont le même caractère, mais on vient de la montrer – avec la remarque que $\chi_i(1) = \dim V_i$.

(Au passage, on note que $\dim V_i > 0$ évidemment, donc V_i apparaît bel et bien comme sous-module de $\mathbb{C}[G]$, ce que nous avions vu dans les exos d'une autre façon.) □

Corollaire 210. *Avec les notations ci-dessus, on a*

$$\sum_i \chi_i(1)^2 = |G|\,,$$

ou encore

$$\sum_i \dim(V_i)^2 = |G|\,.$$

La somme des carrés des dimensions des représentations irréductibles de G est égale à l'ordre du groupe !

Démonstration. Évaluer la relation du théorème en 1. □

Exemple 211. Voyons une utilisation typique du théorème (qui est extrêmement important). Prenons $G = S_3$, et trouvons tous ses caractères irréductibles, puis toutes ses représentations à isomorphisme près. Soit $\chi_1, \ldots, \chi_s$ les caractères irréductibles de S_3. Ce groupe est d'ordre 6, donc

$$\chi_1(1)^1 + \cdots + \chi_s(1)^2 = 6\,,$$

et bien sûr $\chi_i(1)$ est un entier. On voit rapidement qu'il n'y a que deux façons d'écrire 6 comme somme de carrés : ou bien $6 = 1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2$, ou bien $6 = 1^2 + 1^2 + 2^2$. Mais la première option est exclue : en effet si on avait $\chi_i(1) = 1$ pour tout i, les représentations irréductibles de G seraient toutes de dimension 1, et G serait abélien (cf les exos). Ce n'est pas le cas, donc on peut conclure que $s = 3$, avec disons $\chi_1(1) = \chi_2(1) = 1$ et $\chi_3(1) = 2$. On a déjà le nombre de caractères et leurs dimensions !

Par ailleurs, nous connaissons les représentations de dimension 1 : en effet il y a la triviale, appelons-là χ_1, et la signature $\chi_2 \colon S_3 \longrightarrow \{\pm 1\} \subset GL_1(\mathbb{C})$. Rappelons que pour les représentations de dimension 1 il n'y a pas de différence entre la représentation et son caractère.

Voici maintenant la deuxième application vraiment typique du théorème : nous avons tous les caractères sauf un, nous pouvons donc déduire le dernier, en écrivant $\chi_1 + \chi_2 + 2\chi_3 = \chi_{reg}$ et ainsi

$$\chi_3 = \frac{1}{2}\left(\chi_{reg} - \chi_1 - \chi_2\right)\,.$$

On peut lister toutes les valeurs prises par χ_3 ; comme c'est une fonction centrale, il suffit de donner

$$\chi_3(I) = 2\,, \quad \chi_3((12)) = \frac{1}{2}(0-1+1) = 0\,, \quad \chi_3((123)) = \frac{1}{2}(0-1-1) = -1\,.$$

Et voilà le travail. En fait dans les exos nous avons vu que la représentation de S_3 dans $\mathbb{C}^3$ par permutation des coordonnées pouvait s'écrire $\mathbb{C} \oplus V$ en écrivant juste $\mathbb{C}$ pour la représentation triviale, et avec V irréductible de dimension 2. On doit alors avoir $\chi_V = \chi_3$, c'est le seul caractère irréductible de degré 2. Nous avons donc toutes les représentations de S_3, à isomorphisme près.

§5.7 Le nombre de caractères irréductibles

Soit G un groupe fini. Nous allons calculer la dimension de l'espace vectoriel $\mathscr{C}(G)$ de deux façons différentes, et ceci va nous donner le nombre de caractères irréductibles de G, comme on va le voir.

Commençons par :

Lemme 212. *Soit K une classe de conjugaison dans G, et soit $f_K : G \longrightarrow \mathbb{C}$ définie par*

$$f_K(x) = \begin{cases} 1 \text{ si } x \in K\,, \\ 0 \text{ sinon}\,. \end{cases}$$

Alors $f_K \in \mathscr{C}(G)$, et à mesure que K parcourt les classes de conjugaison, la famille des f_K est une base de $\mathscr{C}(G)$.

Démonstration. Il est évident que f_K est centrale. Soient $K_1, \ldots, K_s$ les classes de conjugaison de G, et soit $f \in \mathscr{C}(G)$. Par définition, f est constante sur les classes K_i. Si on prend $x_i \in K_i$, on a la relation

$$f = \sum_i f(x_i) f_{K_i}\,.$$

En effet, si on évalue des deux côtés en x, avec $x \in K_{i_0}$, on obtient $f(x_{i_0})$ des deux côtés. Voilà qui montre que les fonctions $f_{K_1}, \ldots, f_{K_s}$ forment une famille

génératrice de $\mathscr{C}(G)$. C'est aussi une famille libre, puisqu'en présence d'une relation

$$\sum_i \lambda_i f_{K_i} = 0$$

il suffit d'évaluer en x_i pour trouver $\lambda_i = 0$. □

Par contre, il est un peu plus difficile de montrer :

Proposition 213. *Soient $\chi_1, \dots, \chi_s$ les caractères irréductibles de G. Alors $\chi_1, \dots, \chi_s$ est une base de $\mathscr{C}(G)$.*

Démonstration. Montrons d'abord qu'il s'agit là d'une famille libre : si

$$\sum_i \lambda_i \chi_i = 0\,,$$

on applique $(-, \chi_j)$ pour obtenir par le lemme de Schur :

$$\left(\sum_i \lambda_i \chi_i, \chi_j\right) = \sum_i \lambda_i (\chi_i, \chi_j) = \lambda_j = 0\,.$$

Il faut maintenant montrer que cette famille est génératrice. Soit E l'espace vectoriel engendré par les χ_i. On a alors $\mathscr{C}(G) = E \oplus E^\perp$ où $E^\perp$ est l'espace des f tels que $(\chi_i, f) = 0$ pour tout i. En effet, c'est un argument que nous avons donné dans la preuve du corollaire 183. Nous souhaitons donc montrer que $E^\perp = \{0\}$.

Soit $f \in \mathscr{C}(G)$ telle que $(\chi_i, f) = 0$ pour tout i. Prenons une représentation irréductible ρ de G dans l'espace vectoriel V. Introduisons l'endomorphisme $\varphi : V \longrightarrow V$ obtenu en faisant une espèce de moyenne des ρ_x à l'aide de f, plus précisément par la formule

$$\varphi = \sum_{x \in G} \overline{f(x)} \rho_x\,. \tag{*}$$

Alors pour $y \in G$ on a $\rho_y \circ \varphi = \varphi \circ \rho_y$, ou ce qui revient au même, $\rho_y \circ \varphi \circ \rho_y^{-1} = \varphi$, puisque

$$\rho_y \circ \varphi \circ \rho_y^{-1} = \sum_{x \in G} \overline{f(x)} \rho_{yxy^{-1}} = \sum_{z \in G} \overline{f(y^{-1}zy)} \rho_z = \sum_{z \in G} \overline{f(z)} \rho_z = \varphi\,.$$

(La deuxième égalité en posant $z = yxy^{-1}$, la troisième parce que f est centrale.) Mais alors, le lemme de Schur nous dit que $\varphi = \lambda I$, la représentation ρ étant irréductible.

On peut trouver la valeur de λ en prenant la trace : $\mathrm{Trace}(\varphi) = \lambda \dim(V)$, donc $\lambda = \mathrm{Trace}(\varphi)/\dim V$. Mais il est temps d'utiliser le fait que $f \in E^\perp$ (qui n'avait pas encore servi) :

$$\mathrm{Trace}(\varphi) = \sum_{x\in G} \overline{f(x)}\,\mathrm{Trace}(\rho_x) = \sum_{x\in G} \overline{f(x)}\chi_\rho(x) = |G|(f,\chi_\rho) = 0\,,$$

puisque f est supposée orthogonale à tous les caractères irréductibles. D'où $\lambda = 0$, et finalement $\varphi = 0$.

Nous pouvons en déduire que, si ρ est maintenant une représentation *quelconque* de G dans V, et si on définit φ par (*), alors $\varphi = 0$ (il suffit d'écrire V comme une somme d'irréductibles).

Pour conclure, il va nous suffire de prendre pour V la représentation régulière $\mathbb{C}[G]$, avec sa base formée de g^* pour $g \in G$. On calcule simplement

$$\varphi(1^*) = \sum_{x\in G} \overline{f(x)}x^* = 0\,,$$

d'où $\overline{f(x)} = 0$ pour tout $x \in G$, et nous avons enfin montré que $f = 0$. □

Théorème 214. *Le nombre de caractères irréductibles du groupe fini G est égal au nombre des classes de conjugaison dans G.*

Démonstration. D'après le lemme, la dimension de $\mathscr{C}(G)$ est le nombre de classes de conjugaison, et d'après la proposition, c'est le nombre de caractères irréductibles. □

Exemple 215. Soit G un groupe abélien fini d'ordre n. Alors il possède n classes de conjugaisons. D'après le théorème, il possède également n caractères irréductibles $\chi_1, \ldots, \chi_n$. On a la relation

$$\sum_{i=1}^{n} \chi_i(1)^2 = n\,,$$

d'où $\chi_i(1) = 1$ pour tout i. Nous retrouvons le fait que les représentations irréductibles d'un groupe abélien fini sont toutes de dimension 1.

Exemple 216. Prenons $G = S_4$. Il y a 5 classes de conjugaison : ce sont les classes de I, (12), (123), $(12)(34)$ et (1234). Il y a donc 5 caractères irréductibles. Nous connaissons déjà le caractère trivial χ_1 et la signature $\chi_2 : S_4 \longrightarrow \{\pm 1\} \subset GL_1(\mathbb{C})$. Dans les exos, nous avons aussi construit une représentation irréductible de dimension 3, disons ρ. Nous avons aussi introduit l'opération $\otimes$ qui nous permet de considérer $\chi_2 \otimes \rho$, et nous savons qu'elle est également irréductible (par ailleurs, nous allons voir ci-dessous que $\chi_2 \otimes \rho$ n'est pas égale à ρ). Nous avons donc 4 caractères irréductibles, il en manque donc un seul, disons χ_3.

La somme des degrés au carré doit faire $|S_4| = 24$, donc

$$1^2 + 1^2 + 3^2 + 3^2 + \chi_3(1)^2 = 24\,,$$

d'où $\chi_3(1) = 2$. Et enfin, on peut trouver les valeurs du caractère χ_3 à partir des autres, en utilisant la relation

$$\chi_1 + \chi_2 + 2\chi_3 + 3\chi_\rho + 3\chi_2\chi_\rho = \chi_{reg}\,.$$

(On rappelle que le caractère de $\chi_2 \otimes \rho$ est $\chi_2\chi_\rho$.) Pour cela, commençons par remplir un tableau.

	I	(12)	$(12)(34)$	(123)	(1234)
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	2	?	?	?	?
χ_ρ	3	1	-1	0	-1
$\chi_2\chi_\rho$	3	-1	-1	0	1

Pour finir de remplir, dans la colonne (12) par exemple on doit avoir

$$1 - 1 + 2? + 3 - 3 = 0\,,$$

donc ? = 0 ici. (On rappelle que $\chi_{reg}(x) = 0$ pour $x \neq 1$.) Et ainsi de suite.

	I	(12)	$(12)(34)$	(123)	(1234)
χ_1	1	1	1	1	1
χ_2	1	−1	1	1	−1
χ_3	2	0	2	−1	0
χ_ρ	3	1	−1	0	−1
$\chi_2\chi_\rho$	3	−1	−1	0	1

Ce que nous venons d'écrire s'appelle la *table des caractères* du groupe. C'est une matrice carrée, qui n'inclut pas les indications sur les lignes ou sur les colonnes ; c'est-à-dire que la table des caractères de S_4, par exemple, est

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 \\ 2 & 0 & 2 & -1 & 0 \\ 3 & 1 & -1 & 0 & -1 \\ 3 & -1 & -1 & 0 & 1 \end{pmatrix}.$$

Évidemment, cette matrice dépend de la numérotation des classes de conjugaison, et de celle des caractères, donc elle n'est définie que modulo des permutations des lignes et des colonnes.