Chapitre 6

Polynômes

Dans ce chapitre encore, sauf mention explicite du contraire, tous les anneaux et corps considérés sont supposés commutatifs.

6.1 Anneaux de polynômes

6.1.1. *Notations.* Pour tout $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{N}^n$, on note $|\alpha| = \alpha_1 + \ldots + \alpha_n$; on dit que $|\alpha|$ est le *poids* de α. Si A est un anneau, et $x_1, \ldots, x_n \in A$, on note x^α l'élément$x_1^{\alpha_1} \ldots x_n^{\alpha_n}$; on utilisera systématiquement la convention $a^0 = 1$ pour tout $a \in A$, même lorsque $a = 0$. On convient aussi que $\mathbf{N}^0$ est réduit au singleton $\{0\}$.

6.1.2 Proposition. — *Soit A un anneau, $n \in \mathbf{N}$, et considérons le groupe abélien $A^{(\mathbf{N}^n)}$ des familles $(a_\alpha)_{\alpha \in \mathbf{N}^n}$, $a_\alpha \in A$ presque tous nuls. Alors la multiplication*

$$(a_\alpha).(b_\beta) = (c_\gamma) \qquad \text{avec } c_\gamma = \sum_{\alpha+\beta=\gamma} a_\alpha b_\beta$$

définit sur $A^{(\mathbf{N}^n)}$ une structure d'anneau commutatif.

Démonstration. — Il est clair que la multiplication définie ci-dessus est $\mathbf{Z}$-bilinéaire. Alors les deux applications

$$\mu' : ((a_\alpha), (b_\beta), (c_\gamma)) \to ((a_\alpha).(b_\beta)).(c_\gamma) \qquad \mu'' : ((a_\alpha), (b_\beta), (c_\gamma)) \to (a_\alpha).((b_\beta).(c_\gamma))$$

sont $\mathbf{Z}$-trilinéaires ; pour prouver qu'elles sont égales (ce qui exprime l'associativité de la multiplication), il suffit de le faire pour des éléments qui n'ont qu'une composante non nulle. En d'autres termes, si pour tout $\alpha \in \mathbf{N}^n$ on note $\varepsilon_\alpha : A \to A^{(\mathbf{N}^n)}$ l'application qui à $a \in A$ associe la famille dont tous les termes sont nuls sauf peut-etre celui d'indice α qui vaut a, on peut supposer $(a_\alpha) = \varepsilon_\alpha(a)$, $(b_\beta) = \varepsilon_\beta(b)$, $(c_\gamma) = \varepsilon_\gamma(c)$. Il résulte aussitôt de la définition du produit que l'on a $\varepsilon_\alpha(a).\varepsilon_\beta(b) = \varepsilon_{\alpha+\beta}(ab)$. Donc l'égalité $\mu' = \mu''$ se ramène à :

$$\varepsilon_{(\alpha+\beta)+\gamma}((ab)c) = \varepsilon_{\alpha+(\beta+\gamma)}(a(bc))$$

qui est claire puisque les deux membres sont encore égaux à $\varepsilon_{\alpha+\beta+\gamma}(abc)$ à cause de l'associativité de l'addition dans $\mathbf{N}^n$ et de la multiplication dans A. La commutativité de la multiplication est évidente (par un argument analogue, ou directement), et on voit de même que $\varepsilon_0(1)$ est élément neutre pour la multiplication.

6.1.3. *A-algèbres.* Soit A un anneau. Pour les besoins de ce cours, nous appellerons *A-algèbre* tout anneau B contenant A comme sous-anneau (ou contenant un sous-anneau identifié à A, lorsqu'il n'y a pas d'ambigüité sur l'identification.)

Soit B une A-algèbre, $(x_i)_{i\in I}$ une famille (finie ou infinie) d'éléments de B. Nous appellerons combinaison linéaire à coefficients dans A des x_i toute expression de la forme $\sum_{i\in I} a_i x_i$, avec $a_i \in A$ presque tous nuls. Nous dirons que les x_i sont *linéairement indépendants* sur A, si la seule combinaison linéaire nulle des x_i est celle où tous les coefficients sont nuls ; nous dirons que les x_i *engendrent linéairement* B sur A si tout $b \in B$ est combinaison linéaire des x_i. Si les deux conditions sont remplies simultanément, nous dirons que les x_i forment une *base* de B sur A. Alors tout $b \in B$ s'exprime de façon *unique* comme combinaison linéaire des x_i à coefficients dans A.

Soit $x_1, \ldots, x_n$ une famille d'éléments de B, que nous supposerons finie pour simplifier. Nous dirons que les x_i sont algébriquement indépendants sur A, si les x^α, $\alpha \in \mathbf{N}^n$, sont linéairement indépendants sur A ; nous dirons que les x_i *engendrent algébriquement* B sur A, si les x^α engendrent linéairement B sur A. On dit que B est *de type fini* sur A, s'il existe une famille finie d'éléments de B qui engendrent algébriquement B sur A.

Soient B, B' deux A-algèbres. Nous appellerons *homomorphisme de A-algèbres* de B vers B' tout homomorphisme d'anneaux $\varphi : B \to B'$ tel que $\varphi(ax) = a\varphi(x)$ pour tous $x \in B$, $a \in A$. Il revient au même de dire que pour tout $a \in A$ on a $\varphi(a) = a$, puisque la multiplicativité de φ donne alors $\varphi(ax) = \varphi(a)\varphi(x) = a\varphi(x)$ pour tout $x \in B$.

6.1.4 Théorème. — *Soit A un anneau, et $n \in \mathbf{N}$. Comme dans la démonstration de la prop. 6.1.2, on introduit les homomorphismes de groupes abéliens $\varepsilon_\alpha : A \to A^{(\mathbf{N}^n)}$. On note $e_j \in \mathbf{N}^n$, $1 \leq j \leq n$, la base canonique de $\mathbf{Z}^n$. Alors :*

(a) *ε_0 est un homomorphisme d'anneaux, qui permet d'identifier A à un sous-anneau de $A^{(\mathbf{N}^n)}$; nous ferons désormais cette identification.*

(b) *Pour $1 \leq j \leq n$, notons $X_j = \varepsilon_{e_j}(1)$. Alors les X_j sont algébriquement indépendants sur A, et les X^α, $\alpha \in \mathbf{N}^n$, forment une base de $A^{(\mathbf{N}^n)}$ sur A.*

On note désormais $A^{(\mathbf{N}^n)} = A[X_1, \ldots, X_n]$, et on dit que $A[X_1, \ldots, X_n]$ est l'anneau des polynômes en n indéterminées à coefficients dans A. Dans ces notations, on a :

$$\left(\sum_\alpha a_\alpha X^\alpha\right)\left(\sum_\beta b_\beta X^\beta\right) = \sum_\gamma \left(\sum_{\alpha+\beta=\gamma} a_\alpha b_\beta\right) X^\gamma$$

Démonstration. — (a) est évident. Pour *(b)*, il suffit de remarquer que pour tout $\alpha \in \mathbf{N}^n$, $X^\alpha = \varepsilon_\alpha(1)$, et donc il est clair que tout $f \in A^{(\mathbf{N}^n)}$ s'exprime de façon unique comme combinaison linéaire de X^α à coefficients dans A.

6.2 Polynômes et algèbres de type fini

6.2.1. Le résultat suivant, bien que très simple à démontrer, est fondamental. Il exprime que la A-algèbre $A[X_1, \ldots, X_n]$ est "libre" sur les générateurs $X_1, \ldots, X_n$; de même que pour définir une application linéaire entre espaces vectoriels on peut choisir librement les images d'une base, pour définir un homomorphisme de A-algèbres au départ d'une algèbre de polynômes on peut choisir librement les images des générateurs X_j.

Théorème. — (propriété universelle des anneaux de polynômes) *Soit A un anneau, $n \in \mathbf{N}$. Soit B un autre anneau, $\varphi_0 : A \to B$ un homomorphisme, et $\xi_1, \ldots, \xi_n$ des éléments quelconques de B. Alors il exoste un unique homomorphisme d'anneaux $\varphi : A[X_1, \ldots, X_n] \to B$ tel que $\varphi(X_j) = \xi_j$ pour $1 \leq j \leq n$. En particulier, si B est une A-algèbre, il existe un unique homomorphisme de A-algèbres $\varphi : A[X_1, \ldots, X_n] \to B$ tel que $\varphi(X_j) = \xi_j$ pour tout j.*

Démonstration. — L'unicité est claire puisque A et les X_j engendrent $A[X_1, \ldots, X_n]$. Pour l'existence, on écrit :

$$\varphi(\sum_{\alpha} a_\alpha X^\alpha) = \sum_{\alpha} \varphi_0(a_\alpha)\xi^\alpha$$

et on vérifie aussitôt que φ est un homomorphisme d'anneaux possédant les propriétés voulues. Dans le cas où B est elle-même une A-algèbre, on prend simplement $\varphi_0 = \mathrm{Id}_A$, puis on applique le résultat précédent.

6.2.2 Corollaire. — *Soit B une A-algèbre. Alors B est de type fini sur A si et seulement si elle est quotient d'une algèbre de polynômes.*

6.2.3. Réciproquement, il est très commode de partir des algèbres de polynômes lorsqu'on veut *définir* de nouvelles algèbres possédant certaines propriétés. Prenons pour simplifier A égal à un corps commutatif k. Supposons que l'on veuille définir une k-algèbre engendrée par certains éléments $\xi_1, \ldots, \xi_n$ vérifiant certaines identités algébriques, s'exprimant par des formules polynomiales $R_1(\xi_1, \ldots, \xi_n) = 0, \ldots, R_s(\xi_1, \ldots, \xi_n) = 0$, $R_i \in k[X_1, \ldots, X_n]$. Eh bien, c'est immédiat : il suffit de considérer l'algèbre A quotient de $k[X_1, \ldots, X_n]$ par l'idéal engendré par $R_1, \ldots, R_s$, et de prendre pour ξ_j l'image de X_j par la surjection canonique ! Ce quotient est même la solution maximale au problème cherché, en ce sens que toute autre algèbre vérifiant ces conditions est un *quotient* de l'algèbre A ci-dessus. L'exemple le plus classique est celui de la construction de $\mathbf{C}$ comme quotient de $\mathbf{R}[X]$ par l'idéal engendré par $X^2 + 1$, de sorte que l'image ξ de X vérifie $\xi^2 = -1$; si on part d'un corps plus petit comme le corps $\mathbf{Q}$ des nombres rationnels, ce procédé d'adjonction de nouveaux éléments devient extrêmement riche, et est à la base du traitement purement algébrique et exact des calcul faisant intervenir des quantités telles que $\sqrt{2}$, $\sqrt[3]{5} + \sqrt{7}$, et plus généralement des nombres algébriques arbitraires (cf. 6.4.1) par des systèmes de calcul formel tels que `Maple`.

6.2.4. On peut utiliser le thm. 6.2.1 pour donner une description de $A[X_1, \ldots, X_n]$ comme algèbre de polynômes itérée. Soit A' la sous-A-algèbre de $A[X_1, \ldots, X_n]$ engendrée par $X_1, \ldots, X_{n-1}$; alors il est clair que A' s'identifie à $A[X_1, \ldots, X_{n-1}]$; nous

ferons toujours désormais cette identification. En particulier, $A[X_1, \ldots, X_n]$ devient une A'-algèbre.

Proposition. — *L'unique homomorphisme de A'-algèbres $\varphi : A'[X] \to A[X_1, \ldots, X_n]$ appliquant X sur X_n est un isomorphisme.*

Démonstration. — Il est clair que toute $f \in A[X_1, \ldots, X_n]$ possède une unique expression $f = \sum_j f_j X_n^j$, en mettant en facteur les puissances de X_n ; en effet deux polynômes gX_n^i, hX_n^j, $g, h \in A'$, ne peuvent être égaux que si $i = j$, comme on le voit en écrivant g et h dans la base des $X^{\alpha'}$, $\alpha' \in \mathbf{N}^{n-1}$. Ceci exprime précisément le fait que φ est un isomorphisme.

6.2.5 Corollaire. — *Pour tout anneau factoriel A, et pour tout $n \in \mathbf{N}$, l'anneau $A[X_1, \ldots, X_n]$ est factoriel. En particulier, les anneaux $\mathbf{Z}[X_1, \ldots, X_n]$ et $k[X_1, \ldots, X_n]$, où k est un corps, sont factoriels pour tout $n \in \mathbf{N}$.*

Démonstration. — Immédiat par récurrence sur n à partir du thm. 5.5.12.

6.2.6 Corollaire. — *Pour tout anneau nœthérien A, et pour tout $n \in \mathbf{N}$, l'anneau $A[X_1, \ldots, X_n]$ est nœthérien. En particulier, les anneaux $\mathbf{Z}[X_1, \ldots, X_n]$ et $k[X_1, \ldots, X_n]$, oùk est un corps, sont nœthériens pour tout $n \in \mathbf{N}$.*

Démonstration. — Même chose, en utilisant le thm. 5.6.5.

6.2.7 Corollaire. — *Toute algèbre de type fini sur un anneau nœthérien est un anneau nœthérien. En particulier, toute algèbre de type fini sur un corps est un anneau nœthérien.*

Démonstration. — Cela résulte du cor. 6.2.6, du cor. 6.2.2 et de la prop. 5.6.4.

6.2.8. Soit A un anneau, et soit $\mathcal{F}(A^n, A)$ l'anneau de toutes les fonctions sur A^n à valeurs dans A munis des opérations d'addition et de multiplication "point par point". Alors $\mathcal{F}(A^n, A)$ est une A-algèbre si l'on identifie les éléments de A aux fonctions constantes. Soient $x_1, \ldots, x_n \in \mathcal{F}(A^n, A)$ les fonctions coordonnées. Alors l'image de $f = \sum_\alpha a_\alpha X^\alpha \in A[X_1, \ldots, X_n]$ par l'unique homomorphisme de A-algèbres qui pour tout j applique X_j sur la fonction x_j est la fonction

$$(x_1, \ldots, x_n) \to \sum_\alpha a_\alpha x^\alpha \quad \text{noté aussi } f(x_1, \ldots, x_n)$$

On dit que c'est la *fonction polynôme* sur A^n définie par f.

Plus généralement, si B est une A-algèbre, toute $f \in A[X_1, \ldots, X_n]$ peut être considérée comme un élément de $B[X_1, \ldots, X_n]$, et définit donc aussi une fonction polynôme sur B^n, encore notée $x \to f(x)$ dans des notations un peu plus condensées. Pour tout $\xi = (\xi_1, \ldots, \xi_n) \in B^n$, l'application $f \to f(\xi)$ est un homomorphisme de A-algèbres de $A[X_1, \ldots, X_n]$ vers B, appelé *homomorphisme d'évaluation* au point ξ ; c'est d'ailleurs l'unique homomorphisme de A-algèbres qui applique X_j sur ξ_j pour $1 \leq j \leq n$.

Si A est un corps infini K, des considérations élémentaires sur les polynômes prouvent que l'application qui à f associe la fonction polynôme correspondante est injective ;

il n'y a donc pas d'inconvénient dans ce cas à identifier les polynômes aux fonctions correspondantes. Ce résultat s'étend aux anneaux intègres infinis par la considération du corps de fractions.

Il en va tout autrement si A n'est pas intègre, ou si A est fini. Par exemple, si A est le corps $\mathbf{F}_p$ à p éléments, nous verrons que $x^p = x$ pour tout $x \in \mathbf{F}_p$; les polynômes X et X^p définissent donc dans ce cas la même fonction polynôme. En fait, il n'est pas difficile de prouver que pour tout corps fini k, l'application $k[X_1, \ldots, X_n] \to \mathcal{F}(k^n, k)$ qui à f associe sa fonction polynôme est *surjective* ; elle ne peut pas être injective pour la bonne raison que $\mathcal{F}(k^n, k)$ est un ensemble fini, alors que ce n'est jamais le cas pour $k[X_1, \ldots, X_n]$.

6.3 Polynômes homogènes, degré, racines

6.3.1. Soit A un anneau. Pour tout $m \in \mathbf{N}$, on note $A[X_1, \ldots, X_n]_m = \bigoplus_{|\alpha|=m} AX^\alpha$; on dit que les polynômes $f \in A[X_1, \ldots, X_n]_m$ sont *homogènes de degré* m. Clairement toute $f \in A[X_1, \ldots, X_n]$ a une unique écriture $f = \sum_m f_m$, avec f_m homogène de degré m, et les f_m presque tous nuls. On dit que les f_m sont les *composantes homogènes* de f. Notons que si $A = k$ est un corps, ces composantes homogènes sont des k-espaces vectoriels de dimension finie, et si $A = \mathbf{Z}$, des groupes abéliens libres de type fini.

On appelle *degré* (total) de $f \in A[X_1, \ldots, X_n]$, $f \neq 0$, et on note $\deg(f)$, le plus grand entier m tel que $f_m \neq 0$. On pose par convention $\deg(0) = -\infty$.

6.3.2 Proposition. — *Si $f, g \in A[X_1, \ldots, X_n]$ sont homogènes de degré p, q respectivement, fg est homogène de degré $p + q$.*

Démonstration. — Evident.

6.3.3 Proposition. — *Si A est intègre, $A[X_1, \ldots, X_n]$ est intègre pour tout $n \in \mathbf{N}$. Pour tous f, g non nuls dans $A[X_1, \ldots, X_n]$, on a $\deg(fg) = \deg(f)\deg(g)$.*

Démonstration. — Prouvons l'intégrité de $A[X_1, \ldots, X_n]$ récurrence sur $n \geq 1$. Si $n = 1$, on voit comme pour le cas des corps que pour tous $f, g \in A[X]$ non nuls, $\deg(fg) = \deg(f) + \deg(g)$, et en particulier fg est non nul. Dans le cas général, on utilise l'isomorphisme $A[X_1, \ldots, X_n] \simeq A[X_1, \ldots, X_{n-1}][X]$ de la prop. 6.2.4.

Si maintenant $\deg(f) = p$, $\deg(g) = q$, on aura $\deg(fg) \leq p + q$ de manière évidente, et $f_p g_q \neq 0$, d'où l'égalité.

6.3.4. *Polynôme dérivé.* Soit A un anneau. Pour tout polynôme $f = a_0 + a_1 X + \ldots + a_n X^n \in A[X]$ on définit le polynôme dérivé f' par la formule

$$f' = \sum_{j=1}^{n} j a_j X^{j-1}$$

Pour les polynômes à plusieurs variables, on peut définir de manière analogue les polynômes dérivés partiels par rapport à chacune des indéterminées.

On définit par récurrence les polynômes dérivés successifs $f^{(k)}$ par la formule $f^{(0)} = f$, $f^{(k)} = f^{(k-1)\prime}$ si $k > 0$. Cependant, pour éviter les problèmes de division dans les formules de Taylor il est intéressant aussi d'introduire les dérivées réduites $f^{[k]}$. Moralement, on voudrait poser $f^{[k]} = (1/k!)f^{(k)}$. Comme il n'est pas toujours vrai que $k!$ soit inversible dans A (il peut même être nul !) on définit directement $f \to f^{[k]}$ comme étant l'unique application A-linéaire δ_k de $A[X]$ vers elle-même telle que

$$\delta_k(X^j) = \begin{cases} 0 & \text{si } j < k \\ \binom{j}{k} & \text{sinon} \end{cases}$$

Comme $\binom{j}{k}$ est toujours un entier, on a maintenant une définition valable dans *tout* anneau.

6.3.5 Proposition. — (formule de Taylor) *Soit A un anneau. Alors pour tout $f \in A[X]$, avec $\deg(f) = n$, et pour tout $\lambda \in A$ donné, on a*

$$f = \sum_{k=0}^{n} f^{[k]}(\lambda)(X - \lambda)^k$$

Démonstration. — On voit facilement que les $(X-\lambda)^k$, $k \in \mathbf{N}$, forment encore une base de $A[X]$. Si on écrit f dans cette base :

$$f = \sum_{k=0}^{n} a_k(X - \lambda)^k$$

et si on applique δ_k, on voit que $a_k = f^{[k]}(\lambda)$ puisque $\delta_k(X^j)$ est nul pour $j < k$ et $\delta_k(X^k) = 1$.

6.3.6. *Racines.* Soit A un anneau, $f \in A[X]$. On dit que $\lambda \in A$ est une *racine* de f, si $f(\lambda) \neq 0$. Alors on a la proposition suivante :

Corollaire. — *Soit $f \neq 0$ dans $A[X]$, $\lambda \in A$. Alors $f(\lambda) = 0$ si et seulement si il existe $g \in A[X]$ tel que $f = (X - \lambda)g$.*

Démonstration. — La suffisance est claire. La nécessité résulte de la prop. 6.3.5.

6.3.7. On dit que λ est une racine multiple de f, de multiplicité m, s'il existe un polynôme $g \in A[X]$ tel que $f = (X - \lambda)^m g$.

Corollaire. — *Le polynôme f possède une racine multiple d'ordre m en λ si et seulement si $f^{[k]}(\lambda) = 0$ pour $0 \leq k < m$.*

Démonstration. — La suffisance provient de la formule de Taylor. Pour la nécessité, on écrit g dans la base des $(X-\lambda)^j$ comme dans la démonstration de la prop. 6.3.5, ce qui prouve que l'écriture de f dans cette base commence avec un terme en $(X-\lambda)^m$, et on applique δ_k.

6.3.8 Corollaire. — *Si A est intègre, tout $f \in A[X]$ non nul a un nombre de racines dans A au plus égal à son degré.*

6.3.9 Remarque. — Supposons $A \neq \{0\}$ (ce qui est toujours vrai s'il existe $f \neq 0$ dans $A[X]$.) Alors il est facile de voir que pour tout $\lambda \in A$, $X - \lambda$ est simplifiable dans $A[X]$; donc le polynôme g du cor. 6.3.6 est unique ; de plus, $\deg(g) = \deg(f) - 1$. En revanche, pour le cor. 6.3.8 ci-dessus, il est vraiment nécessaire que A soit intègre (on pourra méditer l'exemple de l'équation $X(X-1) = 0$, qui a quatre racines dans $\mathbf{Z}/6\mathbf{Z}$, et huit dans $\mathbf{Z}/30\mathbf{Z}$)

6.4 Eléments algébriques

6.4.1. Le but de cette section 6.4 est de donner quelques résultats élémentaires sur les extensions algébriques, sans prétendre en aucune façon épuiser le sujet. Jusqu'à la fin de 6.4, on fixe un corps k, et un surcorps K de k ; on dit dans cette situation que K est une *extension* de k. Si K_1 et K_2 sont deux extensions de k, les homomorphismes de corps de K_1 vers K_2 induisant l'identité sur k (qui sont aussi les homomorphismes de k-algèbres de K_1 vers K_2) sont souvent appelés *homomorphismes d'extensions* ; rappelons aussi qu'un homomorphisme de corps est toujours injectif (il ne peut pas être nul puisque $\varphi(1) = 1 \neq 0$, et alors $\operatorname{Ker}\varphi = \{0\}$ puisque $\{0\}$ et K_1 sont les seuls idéaux de K_1.)

Pour tout $\xi \in K$, il existe un unique homomorphisme de k-algèbres $\varphi_\xi : k[X] \to K$ tel que $\varphi_\xi(X) = \xi$. Pour tout $f = \sum_j a_j X^j \in k[X]$ on a $\varphi_\xi(f) = \sum_j a_j \xi^j = f(\xi)$. On dit que ξ est *algébrique* sur k, s'il existe $f \in k[X]$ non nul tel que $f(\xi) = 0$, *i.e.*, si φ_ξ n'est pas injective. Dans le cas contraire, on dit que ξ est *transcendant* sur k. Puisque $k[X]$ est un anneau principal, pour tout $\xi \in K$ algébrique sur k il existe un unique $f \in k[\xi]$ unitaire (*i.e.* de coefficient dominant 1) tel que $\operatorname{Ker}\varphi_\xi = (f)$.Comme K est intègre, l'image de $k[X]$ par φ_ξ est un anneau intègre, et donc (f) est un idéal premier, ce qui prouve que f est irréductible ; on dit que f est le *polynôme irréductible* de ξ sur k, et on note $f = \operatorname{Irr}_k(\xi)$.

Pour tous $\xi_1, \ldots, \xi_n \in K$, on note $k(\xi_1, \ldots, \xi_n)$ le plus petit sous-corps de K contenant k et $\xi_1, \ldots, \xi_n$; il est isomorphe au corps de fractions de la sous-k-algèbre de K engendrée par $\xi_1, \ldots, \xi_n$. On dira aussi que $k(\xi_1, \ldots, \xi_n)$ est le sous-corps de K engendré par k et $\xi 1 n$.

6.4.2 Proposition. — *Soit $\xi \in K$. Les conditions suivantes sont équivalentes :*

(i) *ξ est algébrique sur k ;*

(ii) $\dim_k k(\xi) < \infty$.

Si ces conditions sont remplies, $k(\xi)$ est l'ensemble des combinaisons linéaires à coefficients dans k de $1, \xi, \ldots, \xi^{m-1}$, où $m = \deg()\operatorname{Irr}_k(\xi) = \dim_k k(\xi)$; l'entier m est appelé degré *de ξ sur k.*

Démonstration. — Reprenons les notations de 6.4.1. Supposons ξ algébrique sur k, et soit $f = \operatorname{Irr}_k(\xi)$. Alors on a vu que f est irréductible, donc $k[X]/(f)$ est un corps, et il est clair que φ_ξ passe au quotient en un isomorphisme de $k[X]/(f)$ sur $k(\xi)$. Donc $k(\xi)$ est bien de dimension finie sur k, et cette dimension est $m = \deg(f)$.

Réciproquement, si $k(\xi)$ est de dimension finie sur k, φ_ξ ne peut pas être injective, donc ξ est algébrique sur k.

6.4.3 Définition. — On dit que l'extension K est *algébrique*, si tout $\xi \in K$ est algébrique sur k. On dit que l'extension est *finie*, si $\dim_k(K) < \infty$. D'après la prop. 6.4.2, toute extension finie est donc algébrique. Si K est finie, on notera $[K : k]$ la dimension de K sur k.

6.4.4 Proposition. — *Les éléments de K algébriques sur k forment un sous-corps de K.*

Démonstration. — Soit L l'ensemble des éléments de K algébriques sur k. Montrons que L est stable par addition et par multiplication. Soient $\alpha, \beta \in L$, et considérons le corps $k(\alpha, \beta)$ engendré par k et α, β (cf. 6.4.1.) Comme β est algébrique sur k, il est *a fortiori* algébrique sur $k(\alpha)$ (en effet $\mathrm{Irr}_k(\beta)$ peut être considéré comme polynôme à coefficients dans $k(\alpha)$; attention cependant au fait que $\mathrm{Irr}_k(\beta)$ n'est plus nécessairement irréductible sur $k(\alpha)$). D'après la prop. 6.4.2 on a donc $\dim_{k(\alpha)} k(\beta) < \infty$; de même $\dim_k k(\alpha) < \infty$. Mais il est facile de vérifier que si $(e_i)_{1\leq i\leq m}$ est une base de $k(\alpha)$ sur k, et $(f_j)_{1\leq j\leq n}$ une base de $k(\alpha, \beta)$ sur $k(\alpha)$, alors les $e_i f_j$ forment une base de $k(\alpha, \beta)$ sur k. Donc $k(\alpha, \beta)$ est de dimension finie sur k, et on a même la formule

$$[k(\alpha, \beta) : k] = [k(\alpha, \beta) : k(\alpha)][k(\alpha) : k]$$

Toujours d'après la prop. 6.4.2, tous les éléments de $k(\alpha, \beta)$ sont donc algébriques sur k ; c'est en particulier le cas pour $\alpha + \beta$ et $\alpha\beta$. Comme de plus il est clair que L contient k, Lest une sous-k-algèbre de K. De plus, si $\alpha \in L$, $k(\alpha) \subset L$, donc L contient également les inverses de tous ses éléments non nuls, ce qui prouve bien que c'est un corps.

6.4.5 Définition. — On rappelle qu'un corps E est dit *algébriquement clos*, si tout polynôme à coefficients dans E possède une racine dans E. On dit que K est une *clôture algébrique* de k, s'il est à la fois une extension algébrique et un corps algébriquement clos.

6.4.6. Montrons que si l'on sait plonger k dans un corps algébriquement clos, on en a aussi une clôture algébrique :

Proposition. — *Supposons K algébriquement clos. Alors l'ensemble $\overline{k}$ des éléments de K algébriques sur k est une clôture algébrique de k.*

Démonstration. — On sait déjà d'après la prop. 6.4.4 que $\overline{k}$ est un corps ; il est clair que $\overline{k}$ est une ea de k. Montrons que $\overline{k}$ est algébriquement clos. Soit $f = a_0 + a_1 X + \ldots + a_n X^n \in \overline{k}[X]$; comme K est algébriquement clos, f possède une racine ξ dans K. Soit L le corps $k(a_0, \ldots, a_n)$, engendré par k et les a_j. Alors L est une extension finie de k, comme on le voit aussitôt, et ξ est algébrique sur L, donc $L(\xi) = k(a_0, \ldots, a_n, \xi)$ est encore une extension finie de k, ce qui prouve que ξ est algébrique sur k d'après la prop. 6.4.2, et donc que $\xi \in \overline{k}$.

6.4.7 Proposition. — *Supposons K algébrique, et soit E un corps algébriquement clos. Alors tout homomorphisme de corps $\varphi : k \to E$ s'étend à K.*

Démonstration. — Si $[K : k] < \infty$, on raisonne par récurrence sur $m = [K : k]$. Si $m = 1$ on a $K = k$ et il n'y a rien à démontrer. Supposons donc $m > 1$, et soit $\xi \in K$, $\xi \notin k$; soit $f = \mathrm{Irr}_k(\xi)$. Notons $h \to h^\varphi$ l'unique homomorphisme d'anneaux de $k[X]$ vers $E[X]$ égal à φ sur k et appliquant X sur X. Alors si $h = a_0 + a_1X + \ldots + a_nX^n$, on a $h^\varphi = \varphi(a_0) + \varphi(a_1)X + \ldots + \varphi(a_n)X^n$ (en d'autres termes, h^φ est simplement obtenu en appliquant φ à chaque coefficient de h.) Puisque E est algébriquement clos, f^φ possède une racine ξ' dans E. Clairement, le noyau de l'application $h \to h^\varphi(\xi')$ de $k[X]$ vers E est l'idéal (f) ; donc $h \to h^\varphi(\xi')$ passe au quotient en un homomorphisme de corps $\psi : k(\xi) \to E$. Comme $[K : k(\xi)] = m/[k(\xi) : k] < m$, on peut appliquer l'hypothèse de récurrence et conclure que ψ se prolonge à K.

Dans le cas général, on utilise le lemme de Zorn. Soit $\mathcal{F}$ l'ensemble des couples (L, ψ), où L est un sous-corps de K contenant k, et $\psi : L \to E$ un homomorphisme d'extensions. On ordonne $\mathcal{F}$ en posant $(L, \psi) \subset (L', \psi')$ si et seulement si $L \subset L'$ et $\psi'|_L = \psi$. Alors on voit facilement que $\mathcal{F}$ est inductif ; il contient donc un élément maximal (L_0, φ_0). Si $L_0 \neq K$, on prend $\xi \in K$, $\xi \notin L_0$, et on prouve comme ci-dessus que φ_0 s'étend à $L_0(\xi)$, ce qui est contradictoire avec la maximalité de (L_0, φ_0).

6.4.8 Corollaire. — *Soit $\overline{k}$ une clôture algébrique de k. Alors toute extension algébrique de k est isomorphe à un sous-corps de $\overline{k}$ contenant k.*

6.4.9 Théorème. — *Tout corps possède une clôture algébrique, unique à isomorphisme près.*

Démonstration. — Nous admettrons l'existence de la décomposition (voir cependant la remarque ci-après). Prouvons l'unicité. Soient L et L' deux clôtures algébriques de k. D'après la prop. 6.4.7 ci-dessus, il existe un homomorphisme d'extensions $\varphi : L \to L'$, ce qui permet de considérer L comme un sous-corps de L'. Soit maintenant $\xi' \in L'$ quelconque, et soit $f = \mathrm{Irr}_L(\xi')$. Comme L est algébriquement clos, les seuls polynômes irréductibles dans $L[X]$ sont ceux de degré 1 ; donc $f = X - \xi$ pour un certain $\xi \in L$, et $\xi' = \xi \in L$. Mais alors φ est surjective, donc un isomorphisme de L sur L', ce qu'il fallait démontrer.

6.4.10 Remarque. — Bien que nous n'ayons pas démontré l'existence de la clôture algébrique en général, la prop. 6.4.6 permet de l'obtenir chaque fois que l'on sait plonger k dans un corps algébriquement clos ; en particulier, c'est le cas pour tous les corps k que l'on sait plonger dans $\mathbf{C}$. Bien entendu, une condition nécessaire pour cela est que k soit de caractéristique zéro ; ce n'est pas une condition suffisante, mais il n'est pas exagéré de dire que tous les corps de caractéristique zéro que l'on rencontre en théorie des nombres se plongent dans $\mathbf{C}$; de façon plus artificielle, c'est le cas aussi de ceux que l'on rencontre en géométrie algébrique.

En revanche, pour les corps de caractéristique $p > 0$, il faut effectivement faire la construction. C'est déjà bien intéressant dans le cas des corps finis ; malheureusement la construction nous entraînerait un peu loin.

6.4.11 Exemple. — Comme rappelé ci-dessus, la clôture algébrique du corps $\mathbf{Q}$ des nombres rationnels peut se réaliser comme le corps des nombres complexes qui sont algébriques sur $\mathbf{Q}$; on dit simplement que c'est le corps des nombres algébriques, et on le note $\overline{\mathbf{Q}}$. Le degré $[\overline{\mathbf{Q}} : \mathbf{Q}]$ de cette extension est infini ; en effet on voit facilement qu'il existe des polynômes irréductibles dans $\mathbf{Q}[X]$ de degré arbitrairement grand.

Montrons cependant que le corps $\overline{\mathbf{Q}}$ est *dénombrable*. En effet, puisque $\mathbf{Q}$ est dénombrable, tout $\mathbf{Q}$-espace vectoriel de dimension finie est dénombrable, et donc $\mathbf{Q}[X]$ est dénombrable, puisqu'il est réunion d'une suite croissante de $\mathbf{Q}$-espaces vectoriels de dimension finie. *A fortiori*, il n'y a qu'un nombre dénombrable de polynômes irréductibles unitaires à coefficients dans $\mathbf{Q}$. Or chaque $\xi \in \overline{\mathbf{Q}}$ est racine d'un unique tel polynôme f ; on a donc une partition de $\overline{\mathbf{Q}}$ en une famille dénombrable d'ensembles finis, ce qui prouve bien que $\overline{\mathbf{Q}}$ est dénombrable.

6.5 Corps finis

6.5.1. Les notions élémentaires sur les corps exposées dans la section 6.4 suffisent à la description complète des corps finis, comme nous allons le voir maintenant. Ceux-ci constituent des structures algébriques extrêmement remarquables, qui de plus ont trouvé récemment des applications pratiques très importantes dans le domaine des télécommunications notamment (codes correcteurs d'erreurs.)

6.5.2. Le résultat essentiel sur les corps finis est le suivant :

Théorème. — *Pour tout nombre premier p, et tout entier $r \geq 1$, il existe un corps fini $\mathbf{F}_q$ à q éléments, unique à isomorphisme près.*

6.5.3. Bien sûr, pour tout nombre premier p nous connaissons le corps fini $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, et il est clair que $\mathbf{F}_p$ est le seul corps à p éléments : en effet pour tout corps K il existe un unique homomorphisme d'anneaux $\varphi : \mathbf{Z} \to K$ donné par $\varphi(n) = n.1$; dans le cas où K possède p éléments, le groupe additif $(K, +)$ est engendré par 1, donc l'homomorphisme φ est surjectif, et passe au quotient en un isomorphisme $\mathbf{F}_p \to K$.

Plus généralement, si K est fini, φ définit un homomorphisme $\mathbf{F}_p \to K$, où p est la caractéristique de K ; donc tout corps fini peut être vu comme une extension d'un $\mathbf{F}_p$. Une manière d'obtenir l'existence d'un corps fini de cardinal $q = p^r$ est de prouver que pour tout r il existe des polynômes irréductibles sur $\mathbf{F}_p$ de degré r. Il est possible de faire cela directement par un argument de cardinalité ; nous procéderons un peu plus indirectement, mais cette idée jouera quand même un rôle essentiel dans la construction.

6.5.4. Nous avons vu en **??** que pour tout corps K, tout sous-groupe fini de $K^\times$ est cyclique ; en particulier le groupe multiplicatif de tout corps fini K de cardinal q est cyclique de cardinal $q - 1$. On a donc $x^{q-1} = 1$ pour tout $x \in K^\times$; et par conséquent $x^q = x$ pour tout $x \in K$, puisque c'est trivialement vrai pour $x = 0$. Ceci signifie que si p est la caractéristique de K, le polynôme $X^q - X \in \mathbf{F}_p[X]$ a toutes ses racines dans K : il s'écrit $\prod_{\lambda \in K}(X - \lambda)$.

Par conséquent, si un corps de cardinal q existe, les polynômes irréductibles sur $\mathbf{F}_p$ de ses divers éléments sont exactement les facteurs irréductibles dans la factorisation de

$X^q - X$ dans $\mathbf{F}_p[X]$.

6.5.5. Introduisons maintenant un ingrédient essentiel de la discussion, l'homomorphisme de Frobenius :

Proposition. — *Soit K un corps de caractéristique $p > 0$. Alors l'application $F : x \to x^p$ est un homomorphisme de corps de K vers lui-même, appelé homomorphisme de Frobenius. Si K est fini ou algébriquement clos, F est un isomorphisme.*

Démonstration. — Il est clair que $F(xy) = F(x)F(y)$ pour tous $x, y \in K$, et $F(1) = 1$. Pour tous $x, y \in K$ on a :

$$(x+y)^p = \sum_{j=0}^{p} \binom{p}{j} x^j y^{p-j}$$

Or on vérifie facilement que si $0 < j < p$, $\binom{p}{j} \equiv 0 \mod p$. Donc on a bien $(x+y)^p = x^p + y^p$, et F est un homomorphisme de corps. Comme un homomorphisme de corps est toujours injectif, F est injectif. Si K est fini, cela entraîne immédiatement qu'il soit aussi surjectif ; si K est algébriquement clos, la surjectivité provient du fait que pour tout $a \in K$, le polynôme $X^p - a$ a une racine dans K.

6.5.6 Corollaire. — *Pour tout $r \in \mathbf{N}$, l'ensemble des $x \in K$ tels que $x^{p^r} = x$ est un sous-corps de K.*

Démonstration. — Il s'agit de l'ensemble des x tels que $F^r(x) = x$; comme F^r est un homomorphisme de corps pour tout $r \in \mathbf{N}$, l'assertion est immédiate.

6.5.7 Lemme. — *Soit K un corps fini de cardinal q, et soit r un nombre premier. Alors il existe une extension de K de degré r, unique à isomorphisme près.*

Démonstration. — Considérons la factorisation du polynôme $X^{q^r} - X$ dans $K[X]$. Puisque $q = 0 \mod p$, le polynôme dérivé de $X^{q^r} - X$ est le polynôme constant -1, qui ne s'annule jamais. Donc, d'après le cor. 6.3.7 par exemple, $X^{q^r} - X$ n'a que des racines simples dans toute extension de K, ce qui implique que tous ses facteurs irréductibles apparaissent avec multiplicité 1.

Clairement tout $\lambda \in K$ est racine de $X^{q^r} - X$, donc $X^{q^r} - X$ est divisible par $\prod_{\lambda \in K}(X - \lambda)$ Montrons que tous les facteurs irréductibles de degré > 1 de $X^{q^r} - 1$ sont de degré r exactement. Soit f un tel facteur, m son degré, et soit K' le corps $K[X]/(f)$, de sorte que $|K'| = q^m$. Soit α l'image canonique de X dans K'. Ecrivons $q = p^s$, où s est la caractéristique de K, et soit $G = F^s$ la puissance $s^{\text{ième}}$ de l'homomorphisme de Frobenius de K' ; alors l'ensemble des $x \in K'$ tels que $G(x) = x$, qui est aussi l'ensemble des racines dans K' du polynôme $X^q - X$, est égal à K.

Comme f divise $X^{q^r} - X$, on a $\alpha^{q^r} = G^r(\alpha) = \alpha$; et comme α engendre le corps K', on en déduit que $G^r(x) = x$ pour tout x in K', donc que r divise l'ordre de G dans $\mathrm{Aut}\,(K')$. De même, on a $G^m = \mathrm{Id}_{K'}$; si r et m étaient premiers entre eux, on en déduirait que $G = \mathrm{Id}_{K'}$, ce qui est absurde puisque G ne peut pas avoir plus de q points fixes. Donc r divise m ; mais si $m > r$, on a de même une contradiction avec le fait que

G^r ne peut pas avoir plus de q^r points fixes dans K' ; donc finalement $m = r$ comme annoncé.

Ainsi, le corps K' est bien une extension de degré r de K ; et bien sûr le polynôme $X^{q^r} - X$ est scindé sur K'. Soit maintenant L une extension arbitraire de degré r, et montrons que L est isomorphe à K'. Soit $\alpha \in L$, $\alpha \notin K$, et soit $g = \mathrm{Irr}_K(\alpha)$. Comme $\alpha^{q^r} = \alpha$, g est un diviseur irréductible de degré > 1 de $X^{q^r} - X$, donc $\deg(g) = r$, ce qui entraîne que $L = K(\alpha) \simeq K[X]/(g)$. Mais comme $X^{q^r} - X$ est scindé sur K', g a une racine ξ dans K' ; donc $K' = K(\xi) \simeq K[X]/(g)$, ce qui prouve bien que L et K' sont isomorphes.

6.5.8 Remarque. — En réfléchissant un instant sur la démonstration du lemme ci-dessus, on se convainc aussitôt du fait suivant : les polynômes irréductibles de degré r sur K sont exactement les facteurs irréductibles de degré > 1 de $X^{q^r} - 1$. Il y a donc $(q^r - q)/r$ polynômes irréductibles de degré r sur K.

6.5.9. *Démonstration du théorème 6.5.2.* On fixe le nombre premier p et on raisonne par récurrence sur r. Si r est premier, on applique directement le lemme 6.5.7. Sinon, écrivons $r = r's$, avec s premier. Par hypothèse de récurrence, il existe à isomorphisme près un unique corps K de cardinal $p^{r'}$. D'après le lemme 6.5.7, K possède une extension K' de degré s, également unique à isomorphisme près (attention, l'unicité a lieu en tant qu'extension de K, donc on n'a pas encore terminé.) Donc on a déjà l'existence. Soit L un corps arbitraire de cardinal q^r. Alors comme $X^{p^{r'}} - X$ divise $X^{p^r} - X$, le polynôme $X^{p^{r'}} - X$ est scindé sur L, ce qui prouve que si F est l'homomorphisme de Frobenius de L, $F^{r'}$ a $p^{r'}$ racines dans L ; en d'autres termes, L contient un sous-corps de cardinal $p^{r'}$, isomorphe à K par l'hypothèse de récurrence. On peut maintenant appliquer l'assertion d'unicité du lemme 6.5.7, et conclure.

6.6 Fractions rationnelles ; indépendance algébrique

6.6.1. Dans cette section, nous supposerons pour simplifier que l'anneau de base est un corps k. Le corps de fractions de $k[X_1, \ldots, X_n]$, bien défini d'après la prop. 6.3.3 est noté $k(X_1, \ldots, X_n)$, et est appelé corps des fractions rationnelles en n indéterminées sur k.

6.6.2. Soit K un surcorps de k, non nécessairement algébrique. Si K contient des éléments $\xi_1, \ldots, \xi_n$ algébriquement indépendants sur k, l'unique homomorphisme de k-algèbres de $k[X_1, \ldots, X_n]$ vers K appliquant X_j sur ξ_j pour $1 \leq j \leq n$ est *injectif*; d'après la prop. 5.3.10, il se prolonge donc de manière unique en un homomorphisme de corps de $k(X_1, \ldots, X_n)$ vers K.

Ainsi, toute extension non algébrique de k contient un sous-corps isomorphe à $k(X)$. En fait, nous allons voir que la considération des familles algébriquement indépendantes dans K donne une idée de la "taille" de l'extension, et permet de définir un invariant qui joue un rôle important en géométrie algébrique. Nous aurons besoin d'étendre la notion de famille d'éléments algébriquement indépendants au cas d'une famille infinie— mais c'est immédiat : on dit qu'une famille $(\xi_i)_{i \in I}$ d'éléments de K est algébriquement

indépendante sur k, si toute sous-famille finie extraite de (ξ_i) est algébriquement indépendante.

6.6.3 Définition. — Soit K une extension de k. On appelle *base de transcendance* de K sur k, toute famille d'éléments algébriquement indépendants maximale pour l'inclusion.

6.6.4 Proposition. — *Soit K une extension de k. Alors il existe une base de transcendance de K sur k.*

Démonstration. — Exercice (application immédiate du lemme de Zorn.)

6.6.5 Proposition. — *Soit K une extension de k. Alors $(\xi_i)_{i\in I}$ est une base de transcendance de K sur k si et seulement si les ξ_i sont algébriquement indépendants et K est algébrique sur $k(\xi_i)_{i\in I}$.*

Démonstration. — C'est immédiat : si l'extension n'était pas algébrique, il y aurait un $\eta \in K$ transcendant sur $k(\xi_i)_{i\in I}$; mais alors la famille obtenue en rajoutant η aux ξ_i est encore algébriquement indépendante, contredisant la maximalité de la famille (ξ_i).

6.6.6 Théorème. — *Soit K une extension de k. Alors toutes les bases de transcendance de K sur k ont le même cardinal. Ce cardinal est appelé* degré de transcendance *de K sur k.*

Démonstration. — Nous ne ferons la démonstration que dans le cas où K possède une base de transcendance finie ; pour le cas général il faut faire appel à des raisonnements "zorniens" assez lourds et peu instructifs. De plus c'est le cas d'une base de transcendance finie qui apparaît en géométrie algébrique.

Soit donc $(\xi_1, \ldots, \xi_n)$ une base de transcendance de K sur k. On va montrer que pour toute famille $(\eta_1, \ldots, \eta_m)$ d'éléments de K algébriquement indépendants sur k, on a $m \leq n$, et que l'on peut compléter les η_j en une base de transcendance de K par l'adjonction de $n - m$ éléments ξ_i bien choisis.

En fait, montrons que quitte à permuter les ξ_i, $(\eta_1, \xi_2, \ldots, \xi_n)$ est encore une base de transcendance de K sur k. En effet, par maximalité de la famille (ξ_i), il existe $f \in k[Y, X_1, \ldots, X_n]$ non nul tel que $f(\eta_1, \xi_1, \ldots, \xi_n) = 0$. Comme les ξ_i sont algébriquement indépendants, la variable Y doit effectivement apparaître dans f ; et comme η_1 est transcendant sur k, une des X_i doit figurer également ; quitte à permuter les ξ_i, on peut supposer que X_1 figure effectivement dans f. Mais alors on voit que ξ_1 est algébrique sur le corps $k(\eta_1, \xi_2, \ldots, \xi_n)$, et comme K est algébrique sur $k(\xi_1, \ldots, \xi_n)$, il l'est aussi sur $k(\eta_1, \xi_2, \ldots, \xi_n)$, et donc $(\eta_1, \xi_2, \ldots, \xi_n)$ est bien une base de transcendance d'après la proposition 6.6.5 (remarquer que η_1 ne peut pas être algébrique sur $k(\xi_2, \ldots, \xi_n)$ sans quoi ξ_1 le serait aussi.)

Poursuivons la construction : comme η_2 est algébrique sur $k(\eta_1, \xi_2, \ldots, \xi_n)$, il existe $f \in k[Y_1, Y_2, X_2, \ldots, X_n]$ non nulle telle que $f(\eta_1, \eta_2, \xi_2, \ldots, \xi_n) = 0$; comme ci-dessus, on voit que f fait intervenir effectivement Y_2 et au moins un des X_i, $i \geq 2$; quitte à permuter encore les ξ_i, on peut supposer que f fait effectivement intervenir X_2. On conclut de la même façon que $(\eta_1, \eta_2, \xi_3, \ldots, \xi_n)$ est une base de transcendance de K. De proche

en proche, si $m \leq n$, on trouve une permutation des ξ_i telle que $(\eta_1, \ldots, \eta_m, \xi_{m+1}, \ldots, \xi_n)$ soit une base de transcendance de K sur k. Si $m > n$, après n pas on arrive à la conclusion que $(\eta_1, \ldots, \eta_n)$ est une base de transcendance de K ; mais alors η_{m+1} serait algébrique sur $k(\eta_1, \ldots, \eta_n)$, ce qui est absurde. Donc le cas $m > n$ est impossible. En particulier, on voit que si K possède une base de transcendance finie, il ne peut pas y avoir dans K de familles infinies d'éléments algébriquement indépendants, et donc *toute* base de transcendance de K sera finie.

Si $(\eta_1, \ldots, \eta_m)$ est elle-même une base de transcendance de K, on a nécessairement $m = n$; en effet sinon dans la famille $(\eta_1, \ldots, \eta_m, \xi_{m+1}, \ldots, \xi_n)$ trouvée ci-dessus, ξ_{m+1} devrait être algébrique sur $(\eta_1, \ldots, \eta_m)$, ce qui est absurde.

6.6.7. Les premières démonstrations de transcendance pour des nombres complexes "concrets" sont dues à Hermite pour le nombre e, et à Lindemann pour le nombre π ; ce dernier résultat entraîne immédiatement la résolution par la négative du fameux problème de la quadrature du cercle (consistant à trouver une construction par la règle et le compas d'un carré ayant même périmètre qu'un cercle donné) ; on montre en effet facilement que tous les segments que l'on peut construire par la règle et le compas sont de longueur algébrique.

Ce n'est que tout récemment, en 1996, qu'a été démontrée l'indépendance algébrique de e et π. Les résultats concrets d'indépendance algébrique sont encore bien plus difficiles que ceux de transcendance ; pourtant il est très simple de prouver par un argument de cardinalité qu'il doit exister des familles arbitrairement grandes, et même des familles infinies non dénombrables, d'éléments de $\mathbf{C}$ algébriquement indépendants sur $\mathbf{Q}$; l'impression qu'on a d'"avoir triché" vient du fait qu'on a remplacé la question de donner des exemples explicites d'éléments algébriquement indépendants, ou encore plus difficile, de montrer que telle ou telle famille donnée est algébriquement indépendante, par la question beaucoup moins profonde de la simple existence de telles familles.

On a vu dans l'exemple 6.4.11 que la clôture algébrique $\overline{\mathbf{Q}}$ était dénombrable ; comme on sait par ailleurs depuis Cantor que $\mathbf{R}$ n'est pas dénombrable (et donc $\mathbf{C}$ non plus), c'est qu'il doit exister des nombres transcendants (et même, que "la plupart" des nombres complexes sont transcendants). Mais pour tout entier n, l'algèbre $\overline{\mathbf{Q}}[X_1, \ldots, X_n]$ est réunion dénombrable de $\overline{\mathbf{Q}}$-espaces vectoriels de dimension finie, donc est encore dénombrable ; et alors il en va de même de son corps de fractions $\overline{\mathbf{Q}}(X_1, \ldots, X_n)$. Donc si $\mathbf{C}$ possédait une base de transcendance finie, il serait extension algébrique d'un corps dénombrable, et donc lui-même dénombrable (on procède exactement comme pour le cas de $\overline{\mathbf{Q}}$.)

Donc, on voit déjà qu'il existe dans $\mathbf{C}$ des familles finies d'éléments algébriquement indépendants de cardinal aussi grand qu'on veut. Mais en fait, même l'algèbre des polynômes à une infinité *dénombrable* de variables sur un corps dénombrable est dénombrable. Donc, si $\mathbf{C}$ possédait une base de transcendance dénombrable sur $\overline{\mathbf{Q}}$, il serait encore dénombrable, et à nouveau on aurait une contradiction. Vertige ...

6.6.8. C'est cette vision de $\mathbf{C}$ comme clôture algébrique d'un corps de fractions rationnelles en un nombre colossalement infini d'indéterminées qui permet de montrer que quasiment tout corps de caractéristique zéro s'injecte dans $\mathbf{C}$. C'est déjà le cas pour toute

extension de $\mathbf{Q}$ possédant une base de transcendance dont le cardinal ne dépasse pas la puissance du continu ; en particulier pour toute extension de $\mathbf{Q}$ de degré de transcendance fini sur $\mathbf{Q}$ (et plus généralement pour tout corps dénombrable de caractéristique zéro.) Mais en fait, si on supprime un élément ξ dans une base de transcendance de $\mathbf{C}$ sur $\mathbf{Q}$, et qu'on appelle $E \subset \mathbf{C}$ la clôture algébrique du sous-corps de $\mathbf{C}$ engendré par les autres, on voit que E est un corps isomorphe à $\mathbf{C}$, et donc que $\mathbf{C}$ contient un sous-corps $E(\xi)$ isomorphe à $\mathbf{C}(X)$! Donc les choses vont encore beaucoup plus loin : par une généralisation facile, tout corps K qui s'injecte dans une extension de $\mathbf{C}$ dont le degré de transcendance ne dépasse pas la puissance du continu est en fait isomorphe à un sous-corps de $\mathbf{C}$, et on peut utiliser le fait que $\mathbf{C}$ soit algébriquement clos pour prouver l'existence d'une clôture algébrique pour K. Cette fois, on a bien couvert tous les corps de caractéristique zéro qui interviennent non seulement en théorie des nombres, mais aussi en géométrie algébrique.

6.7 Polynômes symétriques

6.7.1. Soit A un anneau non réduit à $\{0\}$. On peut faire agir le groupe symétrique $\mathfrak{S}_n$ sur $A[X_1, \ldots, X_n]$ par automorphismes de A-algèbres, comme suit : à tout $\sigma \in \mathfrak{S}_n$ on associe l'unique automorphisme φ_σ de $A[X_1, \ldots, X_n]$ tel que $\varphi_\sigma(X_j) = X_{\sigma(j)}$. On voit que φ_σ envoie X^α sur $X^{\sigma\alpha}$, où $(\sigma\alpha)_j = \alpha_{\sigma^{-1}(j)}$.

Exemple. — Soit $n = 3$, $A = \mathbf{Z}$, et soit σ la permutation circulaire des indices $1, 2, 3$ donnée par $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$. On a donc $\varphi_\sigma(X_1) = X_2$, $\varphi_\sigma(X_2) = X_3$, $\varphi_\sigma(X_3) = X_1$, et l'extension de φ_σ par multiplicativité donne :

$$\varphi_\sigma(X_1^{\alpha_1} X_2^{\alpha_2} X_3^{\alpha_3}) = \varphi_\sigma(X_1)^{\alpha_1} \varphi_\sigma(X_2)^{\alpha_2} \varphi_\sigma(X_3)^{\alpha_3} = X_2^{\alpha_1} X_3^{\alpha_2} X_1^{\alpha_3} = X^{\sigma\alpha}$$

On voit que les φ_σ respectent les composantes homogènes $A[X_1, \ldots, X_n]_m$.

6.7.2 Théorème. — *L'application $\sigma \to \varphi_\sigma$ définit une action de $\mathfrak{S}_n$ sur $A[X_1, \ldots, X_n]$ par automorphismes de A-algèbres, respectant les degrés.*

Démonstration. — Le théorème affirme que pour tous $\sigma\tau \in \mathfrak{S}_n$ on a $\varphi_{\sigma\tau} = \varphi_\sigma \varphi_\tau$, ce qui est immédiat.

6.7.3 Définition. — On dit que $f \in A[X_1, \ldots, X_n]$ est *symétrique*, si $\varphi_\sigma(f) = f$ pour toute $\sigma \in \mathfrak{S}_n$. On note $A[X_1, \ldots, X_n]^{\mathfrak{S}_n}$ la A-algèbre des polynômes symétriques.

6.7.4 Proposition. — *Soit $f \in A[X_1, \ldots, X_n]$. Alors f est symétrique si et seulement si la fonction $\alpha \to a_\alpha$ qui à $\alpha \in \mathbf{N}^n$ associe le coefficient d'indice α de f est constante sur les orbites de l'action de $\mathfrak{S}_n$ sur $\mathbf{N}^n$.*

Démonstration. — C'est immédiat : à cause de l'unicité de l'expression de f dans la base des X^α, f et $\varphi_\sigma(f)$ doivent avoir les mêmes coefficients. Or si $f = \sum_{\alpha \in \mathbf{N}^n} a_\alpha X^\alpha$:

$$\varphi_\sigma(f) = \sum_\alpha a_\alpha X^{\sigma\alpha} = \sum_\alpha a_{\sigma^{-1}\alpha} X^\alpha$$

donc on doit avoir $a_{\sigma^{-1}\alpha} = a_\alpha$ pour tout $\sigma \in \mathfrak{S}_n$.

6.7.5 Corollaire. — *Pour toute orbite $\mathcal{O}$ de $\mathfrak{S}_n$ dans $\mathbf{N}^n$, posons $e_{\mathcal{O}} = \sum_{\alpha\in\lambda} X^\alpha$. Alors les $e_{\mathcal{O}}$ forment une base de $A[X_1,\dots,X_n]^{\mathfrak{S}_n}$ sur A.*

6.7.6. Les orbites de l'action de $\mathfrak{S}_n$ dans $\mathbf{N}^n$ se rattachent aux partitions d'entiers de la manière suivante : toute $\alpha \in \mathbf{N}^n$ est conjuguée sous l'action de $\mathfrak{S}_n$ à une unique $\lambda = (\lambda_1,\dots,\lambda_n)$ telle que $\lambda_1 \geq \lambda_2 \dots \geq \lambda_n$; en d'autres termes, les orbites correspondant aux polynômes de degré m sont en bijection naturelle avec l'ensemble des partitions de l'entier m en n parts.

Notons $\mathcal{P}$ l'ensemble des $\lambda \in \mathbf{N}^n$ décroissantes, et pour tout $m \in \mathbf{N}$, notons $\mathcal{P}_m = \{\lambda \in \mathcal{P} \mid |\lambda| = m\}$. Notons pour simplifier e_λ l'élément $e_{\mathcal{O}}$ correspondant à l'orbite de λ. On a donc maintenant une base de $A[X_1,\dots,X_n]^{\mathfrak{S}_n}$ et de chaque $A[X_1,\dots,X_n]^{\mathfrak{S}_n}_m$, indexées respectivement par $\mathcal{P}$ et par $\mathcal{P}_m$.

6.7.7 Lemme. — *Ordonnons chaque $\mathcal{P}_m$ par l'ordre lexicographique. Alors pour tous $\lambda \in \mathcal{P}_p$, $\mu \in \mathcal{P}_q$:*

$$e_\lambda e_\mu = e_{\lambda+\mu} + \text{termes d'indice plus petit dans } \mathcal{P}_{p+q}$$

Démonstration. — Notons $\mathcal{O}_\lambda$ l'orbite de λ, $\mathcal{O}_\mu$ celle de μ. Alors

$$e_\lambda e_\mu = \sum_{\alpha\in\mathcal{O}_\lambda,\beta\in\mathcal{O}_\mu} X^\alpha X^\beta = \sum_{\alpha\in\mathcal{O}_\lambda,\beta\in\mathcal{O}_\mu} X^{\alpha+\beta}$$

On remarque que chaque $\lambda \in \mathcal{P}$ est l'unique élément maximal dans son orbite pour l'ordre lexicographique. Montrons que pour tous $\alpha \leq \lambda$, $\beta \leq \mu$, on a $\alpha+\beta \leq \lambda+\mu$, avec égalité si et seulement si $\alpha = \lambda$ et $\beta = \mu$. En effet, pour les premières composantes on a $\alpha_1 \leq \lambda_1$, $\beta_1 \leq \mu_1$, donc $\alpha_1+\beta_1 \leq \lambda_1+\mu_1$, avec égalité si et seulement si $\alpha_1 = \lambda_1$ et $\beta_1 = \mu_1$. Si on a égalité, on poursuit avec les deuxièmes composantes et on trouve de même $\alpha_2+\beta_2 \leq \lambda_2+\mu_2$, avec égalité si et seulement si $\alpha_2 = \lambda_2$ et $\beta_2 = \mu_2$. De proche en proche, on parvient à l'assertion cherchée.

6.7.8 Définition. — Pour $1 \leq p \leq n$, on note σ_p, et on appelle polynôme symétrique élémentaire d'indice p, l'élément e_λ correspondant à $\lambda = (1,\dots,1,0\dots,0)$, avec p composantes égales à 1. En posant $X_I = \prod_{j\in I} X_j$ pour tout $I \subset \{1,\dots,n\}$, on peut donc écrire :

$$\sigma_p = \sum_{\substack{I\subset\{1,\dots,n\}\\|I|=p}} X_I = \sum_{1\leq i_1<\dots<i_p\leq n} X_{i_1}\dots X_{i_p}$$

6.7.9 Théorème. — *Les polynômes symétriques élémentaires constituent une famille de générateurs algébriquement indépendants de $A[X_1,\dots,X_n]^{\mathfrak{S}_n}$; en d'autres termes, toute $f \in A[X_1,\dots,X_n]^{\mathfrak{S}_n}$ s'écrit de manière unique comme polynôme en $\sigma_1,\dots,\sigma_n$.*

Démonstration. — (a) Engendrement. Il suffit de prouver que pour tout $d \in \mathbf{N}$, et toute $\lambda \in \mathcal{P}_d$, e_λ est un polynôme en les σ_j. On procède par récurrence sur d, et pour d fixé, par récurrence sur l'ordre lexicographique des λ. Si $\lambda = 0$, $e_\lambda = 1$ et il n'y a rien à démontrer. Sinon, soit $m \leq n$ le nombre de composantes λ_j non nulles, et soit $\mu = \lambda - \nu_m$, où ν_m est

le n-uplet $(1,\dots,1,0\dots,0)$ avec m composantes égales à 1. Clairement, on a $\mu \in \mathcal{P}_{d-m}$. D'après le lemme 6.7.7 on a

$$e_\lambda = e_\mu \sigma_m - \text{termes d'indice plus petit}$$

donc on conclut par l'hypothèse de récurrence.

(b) Indépendance. On fait une récurrence sur n. Si $n = 0$ il n'y a rien à démontrer. Soit

$$\sum_\alpha \sigma^\alpha = 0 \qquad \text{où } \sigma^\alpha = \sigma_1^{\alpha_1} \dots \sigma_n^{\alpha_n} \tag{$*$}$$

Il s'agit de prouver que tous les a_α sont nuls. Or, si on fait la substitution $X_n = 0$, et si on note $\sigma'_1, \dots, \sigma'_{n-1}$ les polynômes symétriques élémentaires en $X_1, \dots, X_{n-1}$, on obtient : $\sigma_j \to \sigma'_j$ pour $1 \leq j < n$, et $\sigma_n \to 0$. Comme la substitution est un homomorphisme de A-algèbres de $A[X_1, \dots, X_n]$ vers $A[X_1, \dots, X_{n-1}]$, on obtient la relation :

$$\sum_{\alpha_n = 0} a_\alpha \sigma'^\alpha = 0$$

et par indépendance algébrique des σ'_j dans $A[X_1, \dots, X_{n-1}]$, on conclut que $a_\alpha = 0$ lorsque $\alpha_n = 0$. Mais alors, dans la relation $(*)$ on peut mettre σ_n en facteur, et écrire :

$$\sigma_n \left(\sum_\alpha a_\alpha X^{\alpha - e_n} \right) = 0 \qquad \text{où } e_n = (0, \dots, 0, 1)$$

Comme σ_n est toujours un élément simplifiable dans $A[X_1, \dots, X_n]$ (même si A n'est pas intègre !) (puisque la multiplication par σ_n se traduit par une translation au niveau des coefficients), on en déduit que $\sum_\alpha a_\alpha X^{\alpha - e_n} = 0$, et on poursuit ainsi de proche en proche.

6.7.10 Remarque. — On remarque que dans l'écriture $f(X_1, \dots, X_n) = g(\sigma_1, \dots, \sigma_n)$, $f \in A[X_1, \dots, X_n]^{\mathfrak{S}_n}$, le degré total de g n'est autre que le degré de f par rapport à chacune des indéterminées X_j (par symétrie, ce degré est le même pour tous les j.) C'est évident, si on remarque que le degré de chaque σ_m par rapport à chaque X_j est 1.

6.7.11 Remarque. — Le calcul qui permet de passer de l'expression d'un polynôme symétrique dans la base des e_λ à celle dans la base des σ^γ est entièrement indépendant de la nature de l'anneau A ; c'est un calcul de nature combinatoire portant sur les nombres entiers.

Précisons un peu cela. Pour $p \in \{1, \dots, n\}$ notons $\nu_p = (1, \dots, 1, 0 \dots, 0)$ avec p termes égaux à 1, et soit $\gamma \in \mathbf{N}^n$. Si l'on veut calculer lécriture de $\sigma^\gamma = \sigma_1^{\gamma_1} \dots \sigma_n^{\gamma_n}$ dans la base des e_λ, il faut compter pour chaque $\lambda \in \mathcal{P}$ le nombre de manières dont on peut écrire $\lambda = \alpha_1 + \dots + \alpha_d$, $d = |\gamma|$, avec $\alpha_1, \dots, a_{\gamma_1}$ conjugués à ν_1, $\alpha_{\gamma_1+1}, \dots, \alpha_{\gamma_1+\gamma_2}$ conjugués à ν_2, …, $\alpha_{d-\gamma_n+1}, \dots, \alpha_d$ conjugués à ν_n. Les coefficients sont donc des entiers positifs, et par une récurrence facile à partir du lemme 6.7.7, on voit que les λ qui interviennent avec un coefficient non nul vérifient $\lambda \leq \gamma_1 \nu_1 + \dots + \gamma_n \nu_n$.

En fait, l'application $\Lambda : \mathbf{N}^n \to \mathcal{P}$ définie par $\gamma \to \gamma_1\nu_1 + \ldots + \gamma_n\nu_n$ est bijective, d'inverse $\lambda \to \sum_{j=1}^{n-1}(\lambda_j - \lambda_{j+1})\nu_j + \lambda_n\nu_n$. En posant $\mathcal{Q}_m = \{\gamma \in \mathbf{N}^n | \gamma_1 + \ldots + n\gamma_n = m\}$, on voit que Λ se restreint en une bijection de $\mathcal{Q}_m$ sur $\mathcal{P}_m$ pour tout $m \in \mathbf{N}$. Par raison de degré, les seuls e_λ qui peuvent apparaître dans la décomposition de σ^γ avec $\gamma \in \mathcal{Q}_m$ sont les $\lambda \in \mathcal{P}_m$; et toujours d'après le lemme 6.7.7, la matrice qui exprime les σ^γ (ordonnés suivant les valeurs croissantes de $\Lambda(g)$) dans la base des e_λ est triangulaire supérieure avec des 1 sur la diagonale. En notant $N_m = |\mathcal{Q}_m| = |\mathcal{P}_m|$, on a donc une matrice triangulaire $T_m \in \mathbf{M}_{N_m}(\mathbf{Z})$, évidemment inversible. La matrice inverse de T_m, qui est à coefficients entiers, mais non nécessairement positifs, encore avec des 1 sur la diagonale, est la matrice qui donne l'expression des e_λ dans la base des σ^γ. On notera d'ailleurs que ces remarques fournissent une démonstration alternative du thm. 6.7.9.

6.7.12 Exemple. — Traitons par exemple le cas $m = 3$, $n = 3$. Il y a alors trois éléments dans $\mathcal{P}_3$: $(3,0,0) = \Lambda(3,0,0)$, $(2,1,0) = \Lambda(1,1,0)$ et $(1,1,1) = \Lambda(0,0,1)$. On obtient :

$$\begin{aligned}\sigma_3 &= e_{(1,1,1)}\\ \sigma_1\sigma_2 &= e_{(2,1,0)} + 3e_{(1,1,1)}\\ \sigma_1^3 &= e_{(3,0,0)} + 3e_{(2,1,0)} + 6e_{(1,1,1)}\end{aligned}$$

Pour trouver ces formules sans trop de calculs, on remarque que dans $\sigma_1\sigma_2$ il y a neuf termes, et six dans $e_{(2,1,0)}$, donc il y a forcément trois fois $e_{(1,1,1)}$, puisque $(1,1,1)$ est le seul $\lambda < (2,1,0)$ dans $\mathcal{P}_3$, et par ailleurs :

$$\sigma_1^2 = e_{(2,0,0)} + 2e_{(1,1,0)}$$

donne $s_1^3 = s_1 e_{(2,0,0)} + 2s_1 e_{(1,1,0)}$, puis par le même argument que plus haut :

$$s_1 e_{(2,0,0)} = e_{(3,0,0)} + e_{(2,1,0)} \qquad s_1 e_{(1,1,0)} = e_{(2,1,0)} + 3e_{(1,1,1)}$$

En inversant, on trouve :

$$\begin{aligned}e_{(1,1,1)} &= \sigma_3\\ e_{(2,1,0)} &= \sigma_1\sigma_2 - 3\sigma_3\\ e_{(3,0,0)} &= \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3\end{aligned}$$

On remarque qu'en fait ces formules sont valables pour tout $n \geq 3$, en remplaçant bien sûr $(3,0,0)$ par $3\nu_1 = (3,0,\ldots,0)$, $(2,1,1)$ par $\nu_1 + \nu_2$ et $(1,1,1)$ par ν_3.

6.7.13. Voici un exemple un peu plus compliqué. Soit $n = 4$ et cherchons à écrire $e_{(5,3,2,0)}$. On peut mettre σ_4 en facteur si et seulement si tous les λ_j sont > 0 ; ce n'est pas le cas ici. D'après le lemme 6.7.7, $e_{(5,3,2,0)}$ apparaît dans la décomposition de $e^{(4,2,1,0)}.\sigma_3$. Etudions ce produit.

Il s'agit de voir, pour chaque partition μ, de combien de façons on peut écrire $\mu = \alpha + \beta$ avec α conjugué à $(4,2,1,0)$ et β à $(1,1,1,0)$. Vu qu'on ne rajoute que 1 au plus,

si $\alpha_i < \alpha_j$ avec $i < j$ doit avoir $\alpha_j = \alpha_i + 1$ et $\beta_i = 1$, $\beta_j = 0$. Donc ici les α possibles sont $(4,2,1,0)$, $(4,1,2,0)$ et $(4,2,0,1)$. Pour le premier, tous les β sont possibles ; pour $\beta = (1,1,1,0)$, $(1,1,0,1)$, $(1,0,1,1)$, $(0,1,1,1)$ on trouve $(5,3,2,0)$, $(5,3,1,1)$, $(5,2,2,1)$ et $(4,3,2,1)$. Pour chacun des deux autres cas il y a un seul β : pour $\alpha = (4,1,2,0)$ on a $\beta = (1,1,0,1)$ et on trouve $(5,2,2,1)$, pour $\alpha = (4,2,0,1)$ on a $\beta = (1,1,0,1)$ et on trouve $(5,3,1,1)$. On obtient donc :

$$e_{(4,2,1,0)}.\sigma_3 = e_{(5,3,2,0)} + 2e_{(5,3,1,1)} + 2e_{(5,2,2,1)} + e_{4,3,2,1}$$

d'où

$$e_{(5,3,2,0)} = e_{(4,2,1,0)}\sigma_3 - 2e_{(4,2,0,0)}\sigma_4 - 2e_{(4,1,1,0)}\sigma_4 - e_{(3,2,1,0)}\sigma_4$$

Poursuivant ainsi de proche en proche on obtient :

$$\begin{aligned}
e_{(4,2,1,0)} &= e_{(3,1,0,0)}\sigma_3 - 3e_{(3,0,0,0)}\sigma_4 - e_{(2,1,0,0)}\sigma_4 \\
e_{(4,2,0,0)} &= e_{(3,1,0,0)}\sigma_2 - 2e_{(4,1,1,0)} - e_{(3,2,1,0)} - 3e_{(2,0,0,0)}\sigma_4 \\
&= e_{(3,1,0,0)}\sigma_2 - 2e_{(3,0,0,0)}\sigma_3 - e_{(2,1,0,0)}\sigma_3 + 2e_{(2,0,0,0)}\sigma_4 + 2\sigma_2\sigma_4 \\
e_{(4,1,1,0)} &= e_{(3,0,0,0)}\sigma_3 - e_{(2,0,0,0)}\sigma_4 \\
e_{(3,2,1,0)} &= e_{(2,1,0,0)}\sigma_3 - 3e_{(2,0,0,0)}\sigma_4 - 2\sigma_2\sigma_4
\end{aligned}$$

d'où maintenant

$$e_{(5,3,2,0)} = e_{(3,1,0,0)}(\sigma_3^2 - 2\sigma_2\sigma_4) - e_{(3,0,0,0)}\sigma_3\sigma_4 + e_{(2,0,0,0)}\sigma_4^2 - 2\sigma_2\sigma_4^2$$

On trouve encore

$$\begin{aligned}
e_{(3,1,0,0)} &= \sigma_1^2\sigma_2 - 2\sigma_2^2 - \sigma_1\sigma_3 + 4\sigma_4 \\
e_{(3,0,0,0)} &= \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3 \\
e_{(2,0,0,0)} &= \sigma_1^2 - 2\sigma_2
\end{aligned}$$

et finalement :

$$e_{(5,3,2,0)} = \sigma_1^2\sigma_2\sigma_3^2 - 2\sigma_1^2\sigma_2^2\sigma_4 - \sigma_1^3\sigma_3\sigma_4 - 2\sigma_2^2\sigma_3^2 - \sigma_1\sigma_3^3 + 4\sigma_2^3\sigma_4 + \sigma_1^2\sigma_4 + 5\sigma_1\sigma_2\sigma_3\sigma_4 + \sigma_3^2\sigma_4 - 4\sigma_2\sigma_4^2$$

6.7.14 Théorème. — (relations entre coefficients et racines) *Soit k un corps, $f \in k[X]^\bullet$, $f = a_0X^n + a_1X^{n-1} + \ldots + a_n$. Soient $\lambda_1, \ldots, \lambda_n$ les racines de f (dans une clôture algébrique de k.) Alors :*

$$a_j/a_0 = (-1)^j\sigma_j(\lambda_1, \ldots, \lambda_n) \qquad \text{pour } 1 \leq j \leq n$$

Démonstration. — C'est immédiat : il suffit de développer

$$f = a_0\prod_{j=1}^{n}(X - \lambda_j) = a_0\left(X^n + \sum_{m=1}^{n}(-1)^m\left(\sum_{i_1<\ldots<i_m}\lambda_{i_1}\ldots\lambda_{i_m}\right)X^m\right)$$

6.7.15 Corollaire. — *Tout polynôme symétrique en les racines s'exprime comme un polynôme en les a_j/a_0.*

6.7.16 Exemple. — *Discriminant.* Conservons les notations ci-dessus. L'expression

$$\delta(\lambda_1, \ldots, \lambda_n) = \prod_{i<j}(\lambda_i - \lambda_j)$$

s'annule si et seulement si le polynôme P possède une racine multiple ; elle n'est pas symétrique en les λ_j, mais *antisymétrique* : cela signifie que pour tout $\sigma \in \mathfrak{S}_n$, on a $\varphi_\sigma(\delta) = \varepsilon(\sigma)\delta$, où $\varepsilon : \mathfrak{S}_n \to \{\pm 1\}$ est la signature. Donc $D = \delta^2$ est un polynôme symétrique. Il est clair que le degré de D par rapport à chaque λ_j est égal à $2n-2$, puisqu'il y a $n-1$ facteurs $(\lambda_i - \lambda_j)^2$ qui font intervenir un λ_j donné. D'après la remarque rem :degre, D s'exprime donc comme un polynôme de degré $2n-2$ en les a_j/a_0, et $a_0^{2n-2}D$ est un polynôme homogène de degré $2n-2$ en $a_0, \ldots, a_n$, appelé *discriminant* du polynôme f, et noté $\mathrm{Disc}\,(f)$.

Par exemple, si $n = 2$ on a $D = \lambda_1^2 + \lambda_2^2 - 2\lambda_1\lambda_2 = \sigma_1^2 - 4\sigma_2$, d'où $a_0^2 D = a_1^2 - 4a_0a_2$, expression bien connue du discriminant d'une équation du 2e degré. Les calculs pour les degrés plus élevés deviennent vite très compliqués. Par exemple, l'expression du discriminant d'un polynôme de degré 3 est :

$$\mathrm{Disc}\,(a_0X^3 + a_1X^2 + a_2X + a_3) = a_1^2a_2^2 - 4a_0a_2^3 - 4a_1^3a_3 - 27a_0^2a_3^2 + 18a_0a_1a_2a_3$$

Au-delà, il faut faire appel à des systèmes de calcul formel ; il faut bien reconnaître que l'expression explicite est peu éclairante, et l'intérêt du discriminant est surtout théorique. Nous verrons plus loin une expression du discriminant comme un déterminant, qui est plus commode lorsqu'on veut en calculer la valeur pour un polynôme concret.

6.7.17 Exercice. — Voici comment on peut mener le calcul du discriminant d'un polynôme de degré 3 donné ci-dessus. On doit calculer

$$D = (\lambda_1 - \lambda_2)^2(\lambda_1 - \lambda_3)^2(\lambda_2 - \lambda_3)^2$$

qui est un polynôme symétrique homogène de degré total 6, et de degré 4 en chaque λ_j, à coefficients entiers. Dans la décomposition de D dans la base des e_λ ne peuvent donc intervenir que les partitions de 6 en trois parties, dont aucune ne dépasse 4. On voit aussitôt qu'il y a cinq de ces partitions : $(4,2,0)$, $(4,1,1)$, $(3,3,0)$, $(3,2,1)$ et $(2,2,2)$.

En écrivant explicitement les 27 termes du développement de D, ou en calculant le coefficient de certains monômes bien choisis, on arrive alors à l'expression :

$$D = e_{(4,2,0)} - 2e_{(4,1,1)} - 2e_{(3,3,0)} + 2e_{(3,2,1)} - 6e_{(2,2,2)}$$

Il reste maintenant à décomposer les cinq e_λ qui interviennent en polynômes en les σ_j.

On peut utiliser l'algorithme décrit en 6.7.13 et on trouve

$$\begin{aligned}
e_{(4,2,0)} &= \sigma_1^2\sigma_2^2 - 2\sigma_1^3\sigma_3 - 2\sigma_2^3 + 4\sigma_1\sigma_2\sigma_3 - 3\sigma_3^2 \\
e_{(4,1,1)} &= \sigma_1^3\sigma_3 - 3\sigma_1\sigma_2\sigma_3 + 3\sigma_3^2 \\
e_{(3,3,0)} &= \sigma_2^3 - 3\sigma_1\sigma_2\sigma_3 + 3\sigma_3^2 \\
e_{(3,2,1)} &= \sigma_1\sigma_2\sigma_3 - 3\sigma_3^2 \\
e_{(2,2,2)} &= \sigma_3^2
\end{aligned}$$

d'où le résultat en substituant.

6.8 Résultant

6.8.1. Soit k un corps, et soient f et g deux polynômes non nuls dans $k[X]$. On se propose de chercher à quelle condition f et g ne sont pas premiers entre eux, *i.e.* à quelle condition f et g possèdent une racine commune dans une clôture algébrique de k. Nous allons voir que ceci s'exprime par une condition polynomiale sur les coefficients de f et g.

Ecrivons

$$f = a_0X^m + a_1X^{m-1} + \ldots + a_m = a_0\prod_{i=1}^{m}(X - \lambda_i) \qquad \text{avec } a_0 \neq 0$$

$$g = b_0X^n + b_1X^{n-1} + \ldots + b_n = b_0\prod_{j=1}^{n}(X - \mu_j) \qquad \text{avec } b_0 \neq 0$$

(où les λ_i et μ_j sont les racines de f et g respectivement, dans une clôture algébrique $\overline{k}$ de k.) Clairement, f et g ont une racine commune dans $\overline{k}$ si et seulement si

$$a_0^n b_0^m \prod_{i,j}(\mu_j - \lambda_i) = 0$$

Or

$$a_0^n b_0^m \prod_{i,j}(\mu_j - \lambda_i) = b_0^m \prod_{j=1}^{n} f(\mu_j) = (-1)^{mn} a_0^n \prod_{i=1}^{m} g(\lambda_i)$$

Si on voit $\prod_{j=1}^{n} f(\mu_j)$ comme un polynôme en les μ_j à coefficients dans $k[a_0, \ldots, a_m]$, il est clair qu'il est symétrique, et va donc s'exprimer comme polynôme en les b_j/b_0, à coefficients dans $k[a_0, \ldots, a_m]$. Comme on la signalé dans la remarque 6.7.10, le degré total de ce polynôme en les b_j/b_0 est le degré partiel de $\prod_{j=1}^{n} f(\mu_j)$ en μ_1 par exemple, qui est le degré de f. Donc le facteur b_0^m suffit à chasser les dénominateurs, et on aura bien un polynôme en les a_i et les b_j.

Puisque les calculs transformant un polynôme symétrique en polynôme en les polynômes symétriques élémentaires consistent simplement à remplacer les éléments de la base (e_λ) par leur expression en termes des polynômes symétriques élémentaires, et que ces

expressions résultent d'un calcul sur $\mathbf{Z}$, toujours le même et valable dans tout anneau, le polynôme obtenu ci-dessus, vu comme élément de $\mathbf{Z}[a_0, \ldots, a_m, b_0, \ldots, b_n]$, ne dépend pas du corps k dont on est parti. De plus, on obtient le même résultat en partant de $(-1)^{mn} a_0^n \prod_{i=1}^m g(\lambda_i)$ considéré comme polynôme symétrique en les λ_i, puisque les deux polynômes prennent la même valeur si l'on substitue des valeurs rationnelles quelconques pour les a_i et les b_j, avec seulement $a_0 b_0 \neq 0$.

On note dorénavant $R_{m,n} \in \mathbf{Z}[a_0, \ldots, a_m, b_0, \ldots, b_n]$ ce polynôme ; on dit que $R_{m,n}$ est le *résultant* d'ordre (m, n). L'élément de k obtenu en substituant la valeur des coefficients de f et g dans $R_{m,n}$ est appelé résultant de f et g, et noté $\mathrm{Res}\,(f, g)$. On remarque que $\mathrm{Res}\,(f, g)$ est en fait défini pour des polynômes à coefficients dans un anneau A *quelconque* ; en revanche l'interprétation en termes de racines n'a de sens que si A est intègre ; on a alors $\mathrm{Res}\,(f, g) = 0$ si et seulement si f et g ont une racine commune dans la clôture algébrique du corps de fractions de A.

6.8.2. Nous nous proposons maintenant de donner une expression du résultant comme déterminant, qui sera souvent plus commode pour le calcul explicite. Reprenons les polynômes f et g de la section précédente, à coefficients dans un corps k. Soit B l'anneau quotient $k[X]/(g)$, où l'on note (g) l'idéal principal de $k[X]$ engendré par g. Il est clair que $1, X, \ldots, X^{n-1}$ sont linéairement indépendants modulo (g), puisque g ne peut diviser aucun polynôme de degré $< n$, et l'algorithme de division euclidienne montre que tout $h \in k[X]$ s'écrit comme somme d'une combinaison linéaire de $1, X, \ldots, X^{n-1}$ et d'un élément de (g). Donc B est une k-algèbre de dimension finie n comme k-espace vectoriel (attention, si g n'est pas irréductible B n'est pas un corps !).

Soit u l'opérateur de multiplication par $\pi(X)$ dans B, où $\pi : k[X] \to B$ est la surjection canonique. Alors pour tout $h \in k[X]$, l'opérateur $h(u)$ est l'opérateur de multiplication par $h(\pi(X)) = \pi(h)$, donc $h(u) = 0$ si et seulement si $\pi(h) = 0$, *i.e.* si et seulement si g divise h. Ceci montre que le polynôme minimal de u est g ; comme le degré de g est égal à la dimension de B, g est aussi le polynôme caractéristique de u.

Par conséquent, si on trigonalise u sur une clôture algébrique de k, les éléments diagonaux seront les racines μ_j de g ; et alors il est clair que pour tout $h \in k[X]$, $\det h(u) = \prod_{j=1}^n h(\mu_j)$. En particulier, on a $\mathrm{Res}\,(f, g) = b_0^m \det f(u)$.

Considérons maintenant le déterminant $(m + n) \times (m + n)$:

$$\mathrm{Sylv}\,(f, g) = \begin{vmatrix}
a_m & 0 & 0 & \ldots & \ldots & 0 & b_n & 0 & 0 & \ldots & \ldots & 0 \\
a_{m-1} & a_m & 0 & \ldots & \ldots & 0 & b_{n-1} & b_n & 0 & \ldots & \ldots & 0 \\
a_{m-2} & a_{m-1} & a_m & \ldots & \ldots & 0 & b_{n-2} & b_{n-1} & b_n & \ldots & \ldots & 0 \\
\vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots & \vdots & & & \vdots \\
a_0 & a_1 & a_2 & \ldots & \ldots & 0 & b_0 & b_1 & b_2 & \ldots & \ldots & 0 \\
0 & a_0 & a_1 & \ldots & \ldots & a_m & 0 & b_0 & b_1 & \ldots & \ldots & b_n \\
0 & 0 & a_0 & \ldots & \ldots & a_{m-1} & 0 & 0 & b_0 & \ldots & \ldots & b_{n-1} \\
0 & 0 & 0 & \ddots & & a_{m-2} & 0 & 0 & 0 & \ddots & & b_{n-2} \\
\vdots & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots & \vdots & & \ddots & \vdots \\
0 & 0 & 0 & \ldots & \ldots & a_0 & 0 & 0 & 0 & \ldots & \ldots & b_0
\end{vmatrix}$$

appelé *déterminant de Sylvester* de f et g. Notons V le k-espace vectoriel des polynômes de degré $< m+n$, muni de sa base canonique $1, X, \ldots, X^{m+n-1}$. Alors les n premières colonnes de ce déterminant sont les vecteurs $f, Xf, \ldots, X^{n-1}f$; les m dernières sont les vecteurs $g, Xg, \ldots, X^{m-1}g$. La surjection canonique $k[X] \to B$ se restreint en une surjection $V \to B$, envore notée π, dont le noyau est le sous-k-espace vectoriel de V engendré par les $X^i g$. Pour tout vecteur v de V, il existe une unique combinaison linéaire des $X^i g$ qui quand on la soustrait de v donne un polynôme de degré $< n$. Si on fait cette opération sur les n premières colonnes de la matrice, on ne change pas le déterminant, et on obtient une matrice de la forme

$$\begin{vmatrix}
c_{0,0} & \ldots & \ldots & c_{0,n-1} & b_n & 0 & 0 & \ldots & \ldots & 0 \\
\vdots & & & \vdots & b_{n-1} & b_n & 0 & \ldots & \ldots & 0 \\
\vdots & & & \vdots & b_{n-2} & b_{n-1} & b_n & \ldots & \ldots & 0 \\
c_{n-1,0} & \ldots & \ldots & c_{n-1,n-1} & \vdots & \vdots & \vdots & & & \vdots \\
0 & \ldots & \ldots & 0 & b_0 & b_1 & b_2 & \ldots & \ldots & 0 \\
 & & & & 0 & b_0 & b_1 & \ldots & \ldots & b_n \\
 & & & & 0 & 0 & b_0 & \ldots & \ldots & b_{n-1} \\
\vdots & & & \vdots & 0 & 0 & 0 & \ddots & & b_{n-2} \\
 & & & & \vdots & \vdots & \vdots & & & \vdots \\
0 & \ldots & \ldots & 0 & 0 & 0 & 0 & \ldots & \ldots & b_0
\end{vmatrix}$$

Puisque l'on ne change pas $\pi(v)$ en lui retranchant une combinaison linéaire des $X^i g$, on voit en appliquant π que les $c_{i,j}$ sont les coordonnées de $\pi(f), \ldots, \pi(X^{n-1}f)$ dans la base canonique de B, de sorte que le déterminant $|c_{i,j}|$ est le déterminant de $f(u)$ qui nous intéresse. Comme on a maintenant une matrice triangulaire par blocs, avec un bloc $(c_{i,j})$, et l'autre bloc triangulaire d'ordre m avec des b_0 sur la diagonale, on conclut que $\mathrm{Sylv}\,(f,g) = b_0^m \det f(u) = \mathrm{Res}\,(f,g)$, ce qui est bien le résultat escompté.

Pour des polynômes f et g concrètement donnés, cette écriture sous forme de déterminant est certainement la manière la plus efficace de calculer le résultant, en utilisant bien sûr les techniques de pivotage de Gauss.

6.8.3. Montrons maintenant que le discriminant d'un polynôme, dont l'annulation équivaut à l'existence de racines multiples dans la clôture algébrique, s'exprime aussi comme un résultant :

Proposition. — *Soit f in $k[X]$ non nul. Alors*

$$(-1)^{n(n-1)/2} a_0 \mathrm{Disc}\,(f) = \mathrm{Res}\,(f, f')$$

où f' est le polynôme dérivé de f.

Démonstration. — Ecrivons $f = a_0 \prod_{j=1}^{n} (X - \lambda_j)$. Alors par la formule de Leibniz :

$$f' = a_0 \sum_{j=1}^{n} \prod_{i \neq j} (X - \lambda_i)$$

donc pour toute racine λ_j de f, $f'(\lambda_j) = a_0 \prod_{i \neq j}(\lambda_j - \lambda_i)$, et

$$\mathrm{Res}\,(f, f') = (-1)^{n(n-1)} a_0^{n-1} \prod_j f'(\lambda_j) = a_0^{2n-1} \prod_{i \neq j} (\lambda_j - \lambda_i)$$

en remarquant que $(-1)^{n(n-1)}$ vaut toujours 1. Mais nous avons défini le discriminant de f par $\mathrm{Disc}\,(f) = a_0^{2n-2} \prod_{i<j}(\lambda_i - \lambda_j)^2$, d'où le résultat.