Chapitre 9

Polynômes symétriques et résolution des équations

Version du 13 décembre 2004

9.1 Polynômes symétriques

9.1.1 Groupe symétrique et polynômes symétriques

Définition 9.1.1 *On note S_n le groupe des permutations de $\{1,\dots,n\}$, c.-à-d., des bijections de $\{1,\dots,n\}$ sur lui-même. C'est un groupe de cardinal $n!$, car une permutation σ est déterminée par la donnée de $\sigma(1)$, pour lequel il y a n choix, puis de $\sigma(2)$, pour lequel il reste $n-1$ choix, etc.*

Définition 9.1.2 *Soient k un corps et A une k-algèbre. Un k-automorphisme de A est un automorphisme d'anneau $\phi : A \xrightarrow{\sim} A$ tel que $\phi(\lambda) = \lambda$ pour tout $\lambda \in k$. Il est clair que l'ensemble des k-automorphismes de A forme un groupe ; on le note $\mathrm{Aut}_k(A)$.*

Lemme 9.1.1 *Soit k un corps. Tout élément $\sigma \in S_n$ induit un k-automorphisme ϕ_σ de la k-algèbre $k[X_1,\dots,X_n]$, défini par*

$$(*) \qquad \phi_\sigma(X_i) = X_{\sigma(i)}, \quad \forall i = 1,\dots,n.$$

L'application $\sigma \mapsto \phi_\sigma$ est un isomorphisme de S_n sur un sous-groupe du groupe des k-automorphismes de $k[X_1,\dots,X_n]$.

Démonstration. D'après la propriété universelle de $A := k[X_1,\dots,X_n]$, il existe, pour tout $\sigma \in S_n$, un unique morphisme de k-algèbres $\phi_\sigma : A \to A$

vérifiant $(*)$. De plus, il résulte de $(*)$ que $\phi_{\mathrm{id}} = \mathrm{id}_A$ et que $\phi_\sigma \circ \phi_\tau = \phi_{\sigma\tau}$. Ceci entraîne, d'une part, que chaque ϕ_σ est un automorphisme de A, d'inverse $\phi_{\sigma^{-1}}$, et, d'autre part, que l'application $\sigma \mapsto \phi_\sigma$ est un morphisme de groupes de S_n dans $\mathrm{Aut}_k(A)$. Enfin, $(*)$ montre aussi que $\phi_\sigma = \mathrm{id}_A$ ssi $\sigma = \mathrm{id}$, et donc $\sigma \mapsto \phi_\sigma$ est un isomorphisme de S_n sur le sous-groupe $\{\phi_\sigma\}_{\sigma \in S_n}$ de $\mathrm{Aut}_k(A)$. □

Notation Pour tout $P \in k[X_1, \ldots, X_n]$, on écrira simplement $\sigma(P)$ au lieu de $\phi_\sigma(P)$.

Définition 9.1.3 *Soit $P \in k[X_1, \ldots, X_n]$. On dit que P est un* **polynôme symétrique** *si l'on a $\sigma(P) = P$ pour tout $\sigma \in S_n$, c.-à-d., si P est invariant par toute permutation des variables $X_1, \ldots, X_n$. On note*

$$k[X_1, \ldots, X_n]^{S_n}$$

la sous-algèbre des polynômes symétriques. (On voit facilement que c'est une sous-algèbre.)

Exemple 9.1.1 Soit $n = 2$. Les polynômes $X_1 + X_2$, $X_1 X_2$, et $X_1^2 X_2 + X_2^2 X_1$ sont symétriques. Le polynôme $X_1 + X_2^2$ ne l'est pas.

Définition 9.1.4 (Polynômes symétriques élémentaires) *On pose :*

$$\begin{array}{rcl} \sigma_1 & = & X_1 + \cdots + X_n, \\ \sigma_2 & = & \sum_{1 \le i < j \le n} X_i X_j, \\ \vdots & & \\ \sigma_k & = & \sum_{1 \le i_1 < \cdots < i_k \le n} X_{i_1} \cdots X_{i_k}, \\ \vdots & & \\ \sigma_n & = & X_1 \cdots X_n. \end{array}$$

Ce sont des polynômes symétriques, appelés les **polynômes symétriques élémentaires**.

9.1.2 Relations entre coefficients et racines d'un polynôme

Soient k un corps et $P = X^n + a_1 X^{n-1} + \cdots + a_n$ un polynôme unitaire, de degré $n \ge 1$, à coefficients dans k. Soit K une extension de k dans laquelle P est scindé. Alors, dans $K[X]$, on a l'égalité

$$P = (X - x_1)(X - x_2) \cdots (X - x_n), \tag{1}$$

où $x_1, \ldots, x_n$ sont les racines de P dans K, non nécessairement distinctes, c.-à-d., comptées avec leur multiplicité.

Développons le terme de droite de (1). Le coefficient de X^n est, bien sûr, 1. Celui de X^{n-1} est $-(x_1 + \cdots + x_n)$, c.-à-d., $-\sigma_1(x_1, \ldots, x_n)$, et celui de X^{n-2} est $\sum_{i<j} x_i x_j = \sigma_2(x_1, \ldots, x_n)$. Plus généralement, le coefficient de X^{n-r} est

$$(-1)^r \sum_{1 \leq i_1 < \cdots < i_r \leq n} X_{i_1} \cdots X_{i_r} = (-1)^r \sigma_r(x_1, \ldots, x_n).$$

En particulier, le coefficient constant est $(-1)^n \sigma(x_1, \ldots, x_n)$. On a donc obtenu la proposition suivante.

Proposition 9.1.2 (Relation entre coefficients et racines d'un polynôme)
Soient k un corps, $P = X^n + a_1 X^{n-1} + \cdots + a_n$ un polynôme unitaire de degré $n \geq 1$, à coefficients dans k, et $x_1, \ldots, x_n$ les racines de P dans une extension K de k. Pour $i = 1, \ldots, n$, on a

$$a_i = (-1)^i \sigma_i(x_1, \ldots, x_n).$$

9.1.3 Le théorème fondamental des polynômes symétriques

Dans ce paragraphe, k désigne un anneau commutatif arbitraire.

Définition 9.1.5 (Éléments algébriquements indépendants)
Soit A une k-algèbre. On dit que des éléments $e_1, \ldots, e_n \in A$ sont **algébriquement indépendants** *s'ils vérifient la propriété suivante : si $P \in k[X_1, \ldots, X_n]$ et $P(e_1, \ldots, e_n) = 0$, alors $P = 0$. Ceci équivaut à dire que le morphisme de k-algèbres $k[X_1, \ldots, X_n] \to A$ défini par $\phi(X_i) = e_i$ est un isomorphisme ; ceci entraîne, en particulier, que la sous-algèbre de A engendrée par $e_1, \ldots, e_n$ est isomorphe à $k[X_1, \ldots, X_n]$.*

Théorème 9.1.3 (Théorème fondamental des polynômes symétriques)
La sous-algèbre $k[X_1, \ldots, X_n]^{S_n}$ des polynômes symétriques est engendrée sur k par les polynômes symétriques élémentaires $\sigma_1, \ldots, \sigma_n$. De plus, ces éléments sont algébriquement indépendants sur k. Donc, tout polynôme symétrique S s'écrit de façon unique comme un polynôme $P(\sigma_1, \ldots, \sigma_n)$. En résumé, on a un isomorphisme

$$k[X_1, \ldots, X_n]^{S_n} \cong k[\sigma_1, \ldots, \sigma_n],$$

et le terme de droite est un anneau de polynômes.

Exemple 9.1.2 Pour $r \geq 1$, posons $S_r = X_1^r + \cdots + X_n^r$. On a $S_1 = \sigma_1$, et

$$\sigma_1^2 = \sum_{i=1}^{n} X_i^2 + 2\sum_{i<j} X_i X_j,$$

d'où $S_2 = \sigma_1^2 - 2\sigma_2$. De même,

$$\sigma_1^3 = \sum_{i=1}^{n} X_i^3 + 3\sum_{i\neq j} X_i^2 X_j + 3! \sum_{i<j<k} X_i X_j X_k.$$

D'autre part,

$$\sigma_1\sigma_2 = \left(\sum_k X_k\right)\left(\sum_{i<j} X_i X_j\right) = \sum_{i<j}(X_i^2 X_j + X_i X_j^2) + 3\sum_{i<j<k} X_i X_j X_k.$$

On en déduit que $S_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$.

Démonstration. $k[X_1, \ldots, X_n]$ est un k-module libre, de base les monômes $X^\nu := X_1^{\nu_1} \cdots X_n^{\nu_n}$, pour $\nu \in \mathbb{N}^n$. (On rappelle que, dans ce paragraphe, k désigne un anneau commutatif arbitraire.)

Posons $I = \{1, \ldots, n\}$. On regarde $\mathbb{N}^n$ comme l'ensemble des applications $\nu : I \to \mathbb{N}$, $i \mapsto \nu_i = \nu(i)$. On fait agir S_n sur $\mathbb{N}^n$ par la formule :

$$(\sigma\nu)(i) = \nu(\sigma^{-1}(i)), \quad \forall \sigma \in S_n, \nu \in \mathbb{N}^n, i \in I.$$

On vérifie alors que $\sigma(X^\nu) = X^{\sigma(\nu)}$ pour tout ν.

Soit $P \in k[X_1, \ldots, X_n]$ un polynôme symétrique. Écrivons $P = \sum_\nu c_\nu X^\nu$, où les c_ν sont nuls sauf pour un nombre fini d'entre eux. Comme les X^ν sont linéairement indépendants sur k, l'égalité

$$\sum_\nu c_\nu X^\nu = P = \sigma(P) = \sum_\nu c_\nu X^{\sigma(\nu)}$$

entraîne $c_\nu = c_{\sigma(\nu)}$, pour tout $\nu \in \mathbb{N}^n$ et tout $\sigma \in S_n$. Par conséquent, P est combinaison k-linéaire des polynômes symétriques obtenus en additionnant les monômes dans une même orbite :

$$M(\nu) := \sum_{\mu \in S_n \nu} X^\mu.$$

Il est utile, maintenant, de choisir un représentant dans chaque orbite.

Définition 9.1.6 *On dit que $\nu \in \mathbb{N}^n$ est dominant s'il vérifie $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_n \geq 0$. On notera Λ l'ensemble des n-uplets dominants. Il est clair que toute orbite de S_n dans $\mathbb{N}^n$ contient exactement un élément de Λ. Pour $\lambda \in \Lambda$, on désignera par m_λ l'élément considéré plus haut, c.-à-d.,*

$$m_\lambda = \sum_{\mu \in S_n \lambda} X^\mu.$$

Pour démontrer le théorème, on a besoin d'introduire sur $\mathbb{N}^n$ l'ordre lexicographique, défini comme suit.

Définition 9.1.7 *Soient $\mu, \nu \in \mathbb{N}^n$. On dit que $\mu \geq \nu$ si $\mu = \nu$ ou bien s'il existe $i \in \{1, \ldots, n\}$ tel que $\mu_j = \nu_j$ pour $j < i$, et $\mu_i > \nu_i$. On voit facilement que c'est un ordre total, c.-à-d., quelques soient $\mu, \nu \in \mathbb{N}^n$, on a $\mu \leq \nu$ ou $\nu \leq \mu$.*

Remarque 9.1.1 Étant donné $\nu \in \mathbb{N}^n$, l'ensemble des $\mu \leq \nu$ est en général infini. Par exemple, si $n = 2$ et $\nu = (1, 0)$ alors $\nu > (0, i)$, pour tout $i \in \mathbb{N}$. Toutefois, on a le résultat suivant.

Lemme 9.1.4 *Soit $\lambda \in \Lambda$. L'ensemble des $\mu \in \Lambda$ tels que $\mu \leq \lambda$ est fini.*

Démonstration. Les coordonnées d'un tel μ vérifient $0 \leq \mu_i \leq \mu_1 \leq \lambda_1$, donc cet ensemble est fini. □

Soit λ un n-uplet dominant. On voit facilement que λ est l'unique élément maximal, pour l'ordre $\geq$, de l'orbite $S_n\lambda$. C.-à-d., on a :

$$(*) \qquad \forall \lambda \in \Lambda, \forall \mu \in S_n\lambda, \qquad \mu \leq \lambda.$$

Le point crucial dans la démonstration du théorème 9.1.3 est le lemme suivant.

Lemme 9.1.5 (Lemme-clé) *Pour tout $\lambda, \lambda' \in \Lambda$, on a*

$$m_\lambda m_{\lambda'} = m_{\lambda+\lambda'} + \sum_{\substack{\theta \in \Lambda \\ \theta < \lambda+\lambda'}} c_\theta m_\theta.$$

Démonstration. D'une part, il résulte de $(*)$ que

$$(1) \qquad m_\lambda m_{\lambda'} = X^{\lambda+\lambda'} + \sum_{\substack{\mu \in \mathbb{N}^n \\ \mu < \lambda+\lambda'}} c_\mu X^\mu.$$

D'autre part, écrivons

$$m_\lambda m_{\lambda'} = \sum_{\theta\in\Lambda} a_\theta m_\theta, \tag{2}$$

où $a_\theta = 0$ sauf pour un nombre fini d'indices. Posons $E := \{\theta \in \Lambda \mid a_\theta \neq 0\}$; c'est un ensemble fini non-vide. Comme l'ordre lexicographique $\leq$ est un ordre total, E admet un unique élément maximal θ_0. On peut donc écrire :

$$m_\lambda m_{\lambda'} = a_{\theta_0} m_{\theta_0} + \sum_{\substack{\theta\in\Lambda \\ \theta<\theta_0}} a_\theta m_\theta. \tag{3}$$

Alors, d'après $(*)$, le monôme X^{θ_0} n'apparaît que dans m_{θ_0}, et θ_0 est un élément maximal de l'ensemble des $\mu \in \mathbb{N}^n$ tels que X^μ intervienne avec un coefficient non nul dans l'écriture de $m_\lambda m_{\lambda'}$. Comparant avec (1), on obtient que $\theta_0 = \lambda + \lambda'$ et $a_{\theta_0} = 1$. On obtient donc que

$$m_\lambda m_{\lambda'} = m_{\lambda+\lambda'} + \sum_{\substack{\theta\in\Lambda \\ \theta<\lambda+\lambda'}} a_\theta m_\theta.$$

Ceci prouve le lemme. □

On peut maintenant terminer la démonstration du théorème 9.1.3. Comme $\sigma_1, \ldots, \sigma_n$ sont invariants par S_n, la sous-algèbre qu'ils engendrent, notée $k[\sigma]$, est contenue dans la sous-algèbre des invariants. Pour montrer l'inclusion réciproque

$$k[X_1, \ldots, X_n]^{S_n} \subseteq k[\sigma] := k[\sigma_1, \ldots, \sigma_n],$$

il suffit de montrer que m_λ est un polynôme en $\sigma_1, \ldots, \sigma_n$, pour tout $\lambda \in \Lambda$. On va montrer ceci par récurrence sur

$$N(\lambda) := |\{\theta \in \Lambda \mid \theta \leq \lambda\}|.$$

Si $N(\lambda) = 0$, alors $\lambda = 0$ et $m_\lambda = 1$. Pour $i = 1, \ldots, n$, posons

$$\varepsilon_i = (1, \ldots, 1, 0, \ldots, 0),$$

où 1 apparaît i fois, et observons que $m_{\varepsilon_i} = \sigma_i$.

Soit maintenant $N \geq 2$ et supposons le résultat établi pour tout $\theta \in \Lambda$ tel que $N(\theta) < N$. Soit $\lambda \in \Lambda$ tel que $N(\lambda) = N$. Si $\lambda = (d, \ldots, d) = d\varepsilon_n$, alors $m_\lambda = \sigma_n^d$. Sinon, soit i l'unique entier ≥ 1 tel que

$$\lambda_1 = \cdots = \lambda_i > \lambda_{i+1} \geq \cdots \lambda_n.$$

Posons $\lambda' = \lambda - \varepsilon_i$. Alors λ' est dominant, et est $< \lambda$. De plus, d'après le lemme précédent, l'on a

$$m_\lambda = \sigma_i m_{\lambda'} - \sum_{\substack{\theta \in \Lambda \\ \theta < \varepsilon_i + \lambda' = \lambda}} a_\theta m_\theta.$$

Par hypothèse de récurrence, $m_\theta \in k[\sigma]$, pour tout $\theta < \lambda$, y compris $\theta = \lambda'$. L'égalité ci-dessus montre alors que $m_\lambda \in k[\sigma]$ Ceci prouve la 1ère assertion du théorème.

Il reste à voir que $\sigma_1, \ldots, \sigma_n$ sont algébriquement indépendants sur k. Soit $P \in k[X_1, \ldots, X_n]$ non nul. Écrivons

$$P = \sum_{\nu \in \mathbb{N}^n} c_\nu X^\nu,$$

et soit $E = \{\nu \mid c_\nu \neq 0\}$. C'est une ensemble fini non vide. Comme, pour $i = 1, \ldots, n$,

$$\sigma_i = \varepsilon_i = X_1 X_2 \cdots X_i + \text{ monômes plus petits},$$

on déduit du lemme-clé 9.1.5 que, pour tout $\nu \in \mathbb{N}^n$,

$$\sigma_1^{\nu_1} \cdots \sigma_n^{\nu_n} = X_1^{\nu_1 + \cdots + \nu_n} X_2^{a_2 + \cdots + a_n} \cdots X_n^{a_n} + \text{ monômes plus petits}.$$

Ceci conduit à considérer sur $\mathbb{N}^n$ l'ordre $\preceq$ suivant. Observons que l'application $\phi : \mathbb{N}^n \to \mathbb{N}^n$,

$$(\nu_1, \ldots, \nu_n) \mapsto (\sum_{i=1}^n \nu_i, \sum_{i=2}^n \nu_i, \ldots, \nu_n)$$

est injective, car la donnée de $\phi(\nu)$ permet de retrouver ν_n, puis ν_{n-1}, etc. On pose alors

$$\nu \preceq \nu' \Leftrightarrow \phi(\nu) \leq \phi(\nu');$$

c'est une relation d'ordre sur $\mathbb{N}^n$. Soit ν_0 un élément maximal de E pour $\preceq$. Alors $c_{\nu_0} \neq 0$, et $P(\sigma_1, \ldots, \sigma_n)$ égale $c_{\nu_0} X^{\nu_0}$ plus une combinaison linéaire finie de monômes X^μ, avec $\mu \not\succeq \nu_0$ pour l'ordre lexicographique. Par conséquent, $P(\sigma_1, \ldots, \sigma_n) \neq 0$. Ceci achève la preuve du théorème. $\square$

9.2 L'équation générale de degré n

9.2.1 Action d'un groupe sur une algèbre

Soient G un groupe et E un ensemble quelconque. Notons $\mathrm{Bij}(E)$, ou bien $\mathrm{Aut}_{\mathrm{Ens}}(E)$, le groupe des bijections de E de sur E.

Définition 9.2.1 (Action d'un groupe sur un ensemble)
On dit que G agit sur E si l'on s'est donné un morphisme de groupes, pas nécessairement injectif, $\phi : G \to \mathrm{Bij}(E)$. Pour tout $g \in G$, $x \in E$, on écrit $g \cdot x$, ou simplement gx, au lieu de $\phi(g)(x)$.

L'application $G \times E \to E$, $(g, x) \mapsto gx$ s'appelle l'action de G sur E. On voit facilement que la condition que $\phi : G \to \mathrm{Bij}(E)$ soit un morphisme de groupes équivaut aux deux conditions suivantes : pour tout $x \in E$ et $g, g' \in G$,

$$(A) \qquad 1 \cdot x = x \qquad \text{et} \qquad g \cdot (g'x) = (gg') \cdot x.$$

Donc, se donner une action de G sur E équivaut à se donner une application $G \times E \to E$ vérifiant les deux conditions ci-dessus.

Définition 9.2.2 (Action d'un groupe par automorphismes)
Supposons que l'ensemble E soit muni d'une structure algébrique et notons $\mathrm{Aut}(E)$ le sous-groupe de $\mathrm{Bij}(E)$ formé des bijections $E \xrightarrow{\sim} E$ qui préservent la structure algébrique donnée. Alors, on dit que G opère (ou agit) sur E par automorphismes si l'on s'est donné un morphisme de groupes $G \to \mathrm{Aut}(E)$.

Par exemple, si $E = A$ est un k-algèbre, G opère sur A par automorphismes ssi on s'est donné une action $G \times A \to A$ telle que, pour tout $g \in G$ et $a, b \in A$, l'application $\eta_g : a \mapsto g \cdot a$, soit un automorphisme de k-algèbre.

Exemple 9.2.1 On a vu dans le lemme 9.1.1 que S_n agit par automorphismes d'algèbres sur $k[X_1, \ldots, X_n]$.

Lemme 9.2.1 *Soient A un anneau intègre et K son corps des fractions. Tout automorphisme τ de A se prolonge de façon unique en un automorphisme $\widetilde{\tau}$ de K. De plus, l'application $\tau \mapsto \widetilde{\tau}$ est un morphisme injectif de groupes.*

Démonstration. Soit $\tau \in \mathrm{Aut}(A)$ et soient $a, b \in$, avec $b \neq 0$. L'égalité $a = (ab^{-1})b$ dans K montre que toute extension $\widetilde{\tau}$ de τ doit vérifier

$$(*) \qquad \widetilde{\tau}(ab^{-1}) = \tau(a)\tau(b)^{-1}.$$

Réciproquement, on peut définir une application $\widetilde{\tau} : K \to K$ par la formule ci-dessus. Elle est bien définie, car si $ab^{-1} = cd^{-1}$ alors $ad = bc$, d'où $\tau(a)\tau(d) = \tau(b)\tau(c)$.

On vérifie alors sans peine que $\widetilde{\tau}$ est un automorphisme de K. De plus, $(*)$ montre que $\widetilde{\mathrm{id}_A} = \mathrm{id}_K$ et que $\widetilde{\sigma\tau} = \widetilde{\sigma}\widetilde{\tau}$. Donc $\tau \mapsto \widetilde{\tau}$ est un morphisme de groupes, de $\mathrm{Aut}(A)$ vers $\mathrm{Aut}(K)$. Il est de plus injectif, puisque $\widetilde{\tau}(a) = \tau(a)$, pour tout $a \in A$. □

Définition 9.2.3 (Sous-algèbre des invariants)

Soit G un groupe agissant par automorphismes sur une k-algèbre A. On note A^G l'ensemble des éléments invariants, c.-à-d.,

$$A^G = \{a \in A \mid \forall g \in G,\, g \cdot a = a\}.$$

On voit facilement que c'est une sous-algèbre de A.

Proposition 9.2.2 *Soient A un anneau intègre, K son corps des fractions, et G un groupe fini agissant par automorphismes sur A.*

1) *Tout élément de K s'écrit sous la forme a/b, où $b \in A^G$.*

2) *Par conséquent, $K^G = \mathrm{Frac}(A^G)$.*

Démonstration. 1) Soit $x = c/d$ dans K. Comme G est fini, on peut écrire

$$x = \frac{c}{d} = \frac{c\prod_{g\neq 1} g(d)}{\prod_{g\in G} g(d)},$$

et ceci prouve 1). D'autre part, il est clair que

$$\mathrm{Frac}(A^G) = \{ab^{-1} \mid a, b \in A^G,\, b \neq 0\}$$

est un sous-corps de K^G. Réciproquement, soit $x \in K^G$. D'après 1), on peut écrire $x = a/b$, avec $b \in A^G$. Alors, pour tout $g \in G$, l'égalité $g(x) = x$ entraîne $g(a) = a$. Donc $a \in A^G$ et $x \in \mathrm{Frac}(A^G)$. Ceci prouve la proposition. □

9.2.2 Fractions rationnelles symétriques

Soit k un corps et soient $X_1, \ldots, X_n$ des indéterminées. On a vu que le groupe symétrique S_n opère par automorphismes dans $k[X_1, \ldots, X_n]$ et $k(X_1, \ldots, X_n)$. On rappelle que $\sigma_1, \ldots, \sigma_n$ désignent les polynômes symétriques élémentaires.

Théorème 9.2.3 (Théorème des fractions rationnelles symétriques)

1) *On a* $k(X_1, \ldots, X_n)^{S_n} = k(\sigma_1, \ldots, \sigma_n)$.

2) *Par conséquent, l'extension* $k(\sigma_1, \ldots, \sigma_n) \subset k(X_1, \ldots, X_n)$ *est galoisienne, de groupe* S_n.

Démonstration. 1) résulte de la proposition précédente et du théorème fondamental des polynômes symétriques. Le point 2) découle alors du théorème d'Artin. □

Remarque 9.2.1 Soit X une autre indéterminée ; considérons le polynôme suivant, à coefficients dans le corps $k(\sigma_1, \ldots, \sigma_n)$,

$$(*) \qquad Q = X^n - \sigma_1 X^{n-1} + \cdots + (-1)^n \sigma_n.$$

Dans l'extension $k(\sigma_1, \ldots, \sigma_n) \subset k(X_1, \ldots, X_n)$, ce polynôme a pour racines $X_1, \ldots, X_n$, qui sont deux à deux distinctes. Par conséquent, Q est séparable sur le corps $k(\sigma) := k(\sigma_1, \ldots, \sigma_n)$.

9.2.3 L'équation générale de degré n

Soient k un corps, $a_1, \ldots, a_n$ des indéterminées, $K = k(a_1, \ldots, a_n)$ le corps des fractions rationnelles en ces indéterminées. Soit X un autre indéterminées. Considérons le polynôme

$$(**) \qquad P = X^n - a_1 X^{n-1} + \cdots + (-1)^n a_n.$$

L'équation $P(x) = 0$ s'appelle l'équation générale sur k de degré n.

Lorsque $\mathrm{car}(k) = 0$, on voudrait savoir s'il existe une formule "universelle" exprimant les racines de P (dans une extension de K) comme une fonction des a_i obtenue par itération de fonctions polynômiales et de fonctions "extraction de racines" d-èmes (pour tout entier $d \geq 2$). Par exemple, pour $n = 2$, on sait que les racines de

$$X^2 - aX + b = 0$$

sont $(a \pm \sqrt{\Delta})/2$, où Δ désigne le "discriminant" $a^2 - 4b$. On rappelle que cette formule s'obtient en "complétant le carré" de $X - a/2$, c.-à-d., en écrivant

$$X^2 - aX + b = (X - \frac{a}{2})^2 + b - \frac{a^2}{4}.$$

On verra plus loin qu'il existe des formules analogues, mais plus compliquées, pour les équations de degré 3 ou 4, mais qu'il n'existe pas de telles formules pour l'équation générale de degré $n \geq 5$. Commençons par établir le théorème suivant.

Théorème 9.2.4 (S_n est le groupe de Galois de l'équation générale de degré n)
Soit $L = K(x_1, \ldots, x_n)$ un corps de décomposition sur $K = k(a_1, \ldots, a_n)$ du polynôme P ci-dessus. Alors l'extension $K \subset L$ est galoisienne, de groupe S_n. En particulier, les x_i sont deux à deux distincts et P est séparable sur K.

Plus précisément, soient $X_1, \ldots, X_n$ des indéterminées et $\sigma_1, \ldots, \sigma_n$ les polynômes symétriques élémentaires en $X_1, \ldots, X_n$. Alors l'isomorphisme $\phi : k[\sigma] \xrightarrow{\sim} k[a_1, \ldots, a_n]$ défini par $\phi(\sigma_i) = a_i$ pour $i = 1, \ldots, n$ se prolonge en des isomorphismes

$$(\dagger) \qquad \begin{array}{ccc} k(\sigma_1, \ldots, \sigma_n) & \subset & k(X_1, \ldots, X_n) \\ \cong\downarrow & & \cong\downarrow \\ k(a_1, \ldots, a_n) & \subset & k(x_1, \ldots, x_n) \end{array}$$

Démonstration. On a vu que les σ_i sont algébriquement indépendants donc engendrent un anneau de polynômes. Par la propriété universelle, il existe un unique ϕ comme indiqué, et c'est un isomorphisme puisque les a_i sont algébriquement indépendants. Par conséquent, ϕ induit un isomorphisme des corps de fractions, qu'on désignera encore par ϕ.

Posons $Q = X^n - \sigma_1 X^{n-1} + \cdots + (-1)\sigma_n$. Alors, $k(X_i)$ est un corps de décomposition sur $k(\sigma_i)$ de Q. De plus, $\phi(Q) = P$ et, par hypothèse, $L = K(x_i)$ est un corps de décomposition de P sur K. Donc, par le 1er théorème fondamental 7.3.5, ϕ se prolonge en un isomorphisme $\psi : k(X_i) \xrightarrow{\sim} L$.

De plus, ψ induit une bijection entre l'ensemble des racines de Q et de P. Par conséquent, les x_i sont deux à deux distincts et P est séparable sur K. Donc, d'après le 2ème théorème fondamental 7.4.4, l'extension $K \subset L$ est galoisienne. Déterminons son groupe de Galois $\mathrm{Gal}(L/K) = \mathrm{Aut}_K(L)$.

On voit facilement que l'application

$$\tau \mapsto \psi \circ \tau \circ \psi^{-1}$$

est un isomorphisme de $G = \mathrm{Aut}_{k(\sigma_i)}(k(X_i))$, dont l'isomorphisme inverse est donné par $\sigma \mapsto \psi^{-1} \circ \sigma \circ \psi$. On obtient donc, en utilisant le théorème 9.2.3, les isomorphismes

$$\mathrm{Gal}(L/K) \cong \mathrm{Gal}(k(X_1, \ldots, X_n)/k(\sigma_1, \ldots, \sigma_n)) \cong S_n.$$

Ceci prouve le théorème. □

Remarque 9.2.2 Signalons également une autre façon de montrer que l'extension $K \subset L$ est galoisienne, sans utiliser le 2ème théorème fondamental 7.4.4. Posant $H = \mathrm{Aut}_K(L)$, il suffit de montrer que $L^H = K$. Or, comme $L = \psi(k(X_i))$ et $H = \psi G \psi^{-1}$, l'on a

$$L^H = \{\psi(x) \mid \forall \tau \in G,\, \psi(x) = \psi(\tau(x))\},$$

d'où

$$L^H = \psi(k(X_i)^G) = \psi(k(\sigma_i)) = K.$$

9.3 Le corps $\mathbb{C}$ est algébriquement clos

Il existe de nombreuses démonstrations du fait que $\mathbb{C}$ est algébriquement clos. Nous allons en donner deux. La 1ère n'utilise que des méthodes élémentaires d'analyse ; elle est attribuée à Argand, en 1814 (voir [Esc, p.5]). La 2ème est une jolie application de la théorie de Galois ; elle nécessite quelques préliminaires sur les groupes que l'on développera plus bas.

9.3.1 La démonstration d'Argand

Soit $P \in \mathbb{C}[X]$ un polynôme de degré $n \geq 1$. Sans perte de généralité, on peut supposer P unitaire. Écrivons

$$P = X^n + a_1 X^{n-1} + \cdots + a_n.$$

Raisonnons par l'absurde et supposons que P ne s'annule pas sur $\mathbb{C}$. Alors, en particulier, $a_n \neq 0$. Notons $|\cdot|$ la norme usuelle sur $\mathbb{C}$, c.-à-d., si $z = x+iy$ alors

$$|z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}.$$

Comme $\lim_{|z|\to+\infty} |P(z)| = +\infty$, il existe $R > 0$ tel que

$$|z| \geq R \Rightarrow |P(z)| > |a_n|.$$

Explicitement, on peut prendre $R = \max\{1, 2na\}$, où $a = \max_{i=1}^n |a_i|$. En effet, pour $z \geq R$ et $d = 1, \ldots, n$, on a $|z^d| \geq |z| \geq 2na$ d'où

$$\left|\sum_{d=1}^{n} \frac{a_d}{z^d}\right| \leq \sum_{d=1}^{n} \frac{|a_d|}{2na} \leq \frac{1}{2}.$$

Comme $|u - v| \geq |u| - |v|$, on obtient que, pour $|z| \geq R$, on a

$$|P(z)| = |z^n| \cdot \left|1 - \sum_{d=1}^{n} \frac{a_d}{z^d}\right| \geq |z| \cdot \frac{1}{2} \geq n|a_n|.$$

Comme le disque D de centre 0 et de rayon R est compact, la fonction continue $f : z \mapsto |P(z)|$ y atteint son minimum r_0, et $r_0 > 0$ puisqu'on a supposé que P ne s'annule pas. Comme de plus

$$r_0 \leq |P(0)| = |a_n| \leq f(z), \quad \forall z \notin D,$$

alors r_0 est le minimum de f sur $\mathbb{C}$ tout entier. Soit $z_0 \in D$ tel que $f(z_0) = r_0$. En remplaçant z par $z - z_0$ et $P(z)$ par $Q(z) := f(z_0)^{-1}P(z - z_0)$, on se ramène au cas où $z_0 = 0$ et où $Q(0) = 1$ est le minimum de $g = |Q|$ sur $\mathbb{C}$.

Observons que Q est, comme P, de degré n. Notons k l'ordre d'annulation en 0 de $Q - 1$. On peut alors écrire

$$Q(X) = 1 + b_k X^k + \cdots + b_n X^n.$$

avec $b_k b_n \neq 0$. Écrivons $b_k = re^{i\theta}$, avec $r > 0$ et $\theta \in [0, 2\pi[$. Soit r' l'unique racine k-ème de r dans $\mathbb{R}_+^*$. Pour $\varepsilon \in \mathbb{R}_+^*$, posons

$$z_\varepsilon = \frac{\varepsilon}{r'} e^{i(\pi-\theta)/k}, \quad \text{et} \quad q(\varepsilon) = Q(z_\varepsilon).$$

Comme $e^{i\pi} = -1$, alors

$$q(\varepsilon) = 1 - \varepsilon^k + \varepsilon^k h(\varepsilon),$$

où $h(\varepsilon) = \sum_{j=1}^n \frac{b_j}{b_k} z_\varepsilon^j$. Comme $\lim_{\varepsilon \to 0} h(\varepsilon) = 0$, il existe $\varepsilon_0 \in]0, 1[$ tel que

$$\forall \varepsilon \leq \varepsilon_0, \qquad |h(\varepsilon)| \leq \frac{1}{2}.$$

On a alors

$$|Q(z_{\varepsilon_0})| = |1 - \varepsilon_0^k + \varepsilon_0^k h(\varepsilon_0)| \leq 1 - \varepsilon_0^k + \frac{1}{2}\varepsilon_0^k = 1 - \frac{1}{2}\varepsilon_0^k.$$

Ceci contredit l'hypothèse que $1 = Q(0)$ était le minimum de $g = |Q|$ sur $\mathbb{C}$. Cette contradiction montre que l'hypothèse que P ne s'annule pas sur $\mathbb{C}$ est impossible. On a donc démontré le

Théorème 9.3.1 $\mathbb{C}$ *est algébriquement clos.*

9.3.2 Interlude sur les groupes finis : théorème de Sylow et p-groupes

On va donner dans le paragraphe suivant une démonstration du théorème précédent basée sur la théorie de Galois. Pour cela, on aura besoin de deux théorèmes importants sur les groupes finis, que nous énonçons ci-dessous.

Théorème 9.3.2 (Théorème de Sylow)
Soient G un groupe fini, p un nombre premier divisant $|G|$, et p^n la plus grande puissance de p divisant $|G|$. Notons $\mathcal{S}_p(G)$ l'ensemble des sous-groupes de G de cardinal p^n. Alors :

1) *$\mathcal{S}_p(G)$ est non-vide, c.-à-d., il existe au moins un sous-groupe de G de cardinal p^n.*

2) *Les éléments de $\mathcal{S}_p(G)$ sont tous conjugués, c.-à-d., pour tout $H, H' \in \mathcal{S}_p(G)$, il existe $g \in G$ tel que $H' = gHg^{-1}$.*

3) *$|\mathcal{S}_p(G)|$ divise $|G|$ et est congru à 1 modulo p.*

Définition 9.3.1 *Tout sous-groupe de G de cardinal p^n est appelé un* **p-sous-groupe de Sylow de G.**

Remarque 9.3.1 Dans la littérature, les points 1)–3) du théorème sont aussi appelés les trois théorèmes de Sylow.

Démonstration. Pour la démonstration, on renvoie pour le moment à [Pe1, § I.5], ou [BR, Chap. 4], ou [Se, § 8.4], ou [Ja1, § 1.13]. Si le temps le permet, on donnera une démonstration dans la suite du cours. □

Définition 9.3.2 (p-groupes finis) *Soit p un nombre premier. Un groupe fini G est un <u>p-groupe</u> si $|G|$ est une puissance de p.*

Définition 9.3.3 (Centre d'un groupe) *Soit G un groupe. On appelle <u>centre de G</u>, et l'on note $Z(G)$, le sous-ensemble*

$$Z(G) = \{h \in G \mid \forall g \in G,\, hg = gh\}.$$

Lemme 9.3.3 *$Z := Z(G)$ est un sous-groupe de G, tel que $\phi(Z) = Z$ pour tout automorphisme de G. En particulier, $Z(G)$ est un sous-groupe distingué.*

Démonstration. D'abord, $Z(G)$ contient l'élément 1. Soient $z, z' \in Z(G)$ et $g \in G$. D'une part, l'égalité $gz = zg$ entraîne $z^{-1}g = gz^{-1}$. D'autre part, on a $gzz' = zgz' = zz'g$. Ceci montre que $Z(G)$ est un sous-groupe de G. Observons aussi que

$$Z(G) = \{z \in G \mid \forall g \in G, \quad gzg^{-1} = z\}.$$

Soit ϕ un automorphisme de G. Pour tout $g \in G$, on a

$$\phi(z) = \phi(g)\phi(z)\phi(g)^{-1},$$

et comme $\phi(g)$ parcourt G, ceci montre que $\phi(z) \in Z$, c.-à-d., $\phi(Z) \subseteq Z$. De même, $\phi^{-1}(Z) \subseteq Z$ et donc $\phi(Z) = Z$. Ceci prouve le lemme. □

Remarque 9.3.2 1) G est abélien ssi $G = Z(G)$.

2) Pour un groupe $G \neq \{1\}$ arbitraire, on peut avoir $Z(G) = \{1\}$. C'est le cas, par exemple pour $G = S_3$ (exercice !).

Théorème 9.3.4 (Propriétés des p-groupes finis)
Soit G un p-groupe fini, de cardinal p^n, où $n \geq 1$.

1) *On a $Z(G) \neq \{1\}$.*

2) *G contient au moins un sous-groupe distingué de cardinal p^{n-1}.*

Démonstration. Voir [BR, Chap. 4], Théorème 4.6 et Exercice 40, ou [Se, § 8.3], Théorème 14. □

9.3.3 Une démonstration via la théorie de Galois

La démonstration qui suit est tirée de [BR], Appendice, p.287. Pour une autre démonstration, voir [Sa], Appendice au Chap. II. Commençons par rappeler le fait suivant.

Lemme 9.3.5 1) *Tout nombre complexe $z \neq 0$ admet n racines n-ièmes dans $\mathbb{C}$.*

2) *Tout $P \in \mathbb{C}[X]$ de degré 2 est scindé.*

3) *Tout $P \in \mathbb{R}[X]$ de degré impair admet au moins une racine dans $\mathbb{R}$.*

Démonstration. 1) (Ce fait a déjà été utilisé, de façon cruciale, dans la démonstration d'Argand). Posons $z = re^{i\theta}$, avec $r > 0$ et $\theta \in [0, 2\pi[$, et soit $\sqrt[n]{r}$ la racine n-ième de r dans $\mathbb{R}_+$. Alors, les racines n-èmes de z sont les nombres complexes

$$\sqrt[n]{r}\, e^{i\frac{\theta+2k\pi}{n}}, \qquad k = 0, \ldots, n-1.$$

2) On peut supposer P unitaire. Écrivant

$$P = X^2 - 2aX + b = (X-a)^2 + b - 4a^2,$$

on voit que les deux racines de P sont $a \pm \sqrt{b - 4a^2}$.

3) La fonction $\mathbb{R} \to \mathbb{R}$, $x \mapsto P(x)$ est continue. De plus, comme P est de degré impair, on a $\lim_{x\to\pm\infty} P(x) = \pm\infty$. Donc, d'après le théorème des valeurs intermédiaires, il existe $x_0 \in \mathbb{R}$ tel que $P(x_0) = 0$. □

Démontrons maintenant que $\mathbb{C}$ est algébriquement clos. On rappelle qu'en caractéristique 0, tout polynôme est séparable, (Corollaire 7.6.3).

D'abord, l'extension $\mathbb{R} \subset \mathbb{C}$ est de degré 2 et galoisienne, car $\mathbb{C}$ est le corps de décomposition du polynôme X^2+1. Le groupe de Galois $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ est d'ordre 2, donc isomorphe à $\mathbb{Z}/2\mathbb{Z}$; il est engendré par la conjugaison complexe $\tau : z \mapsto \overline{z}$, où $\overline{z} = x - iy$ si $z = x + iy$, $x, y \in \mathbb{R}$.

Soit $P \in \mathbb{C}[X]$ un polynôme irréductible et soit K un corps de décomposition de P sur $\mathbb{C}$. D'après le théorème 7.4.4, l'extension $\mathbb{C} \subseteq K$ est galoisienne. Posons $G_1 = \mathrm{Gal}(L/\mathbb{C})$ et $n = [K : \mathbb{C}] = |G_1|$, et écrivons $n = 2^d r$, avec r impair. Alors,

$$[K : \mathbb{R}] = [K : \mathbb{C}]\,[\mathbb{C} : \mathbb{R}] = 2^{d+1} r.$$

D'après le théorème 7.3.5, l'automorphisme τ de $\mathbb{C}$ se prolonge en un automorphisme $\widetilde{\tau}$ de K. Posons $G_2 = \mathrm{Aut}_{\mathbb{R}}(K)$.

Lemme 9.3.6 *L'extension $\mathbb{R} \subset K$ est galoisienne, de groupe G_2.*

Démonstration. En effet, comme G_2 contient G_1 et $\widetilde{\tau}$, on a

$$K^{G_2} \subseteq K^{G_1} \cap K^{\widetilde{\tau}} = \mathbb{C}^{\tau} = \mathbb{R}.$$

□

Maintenant, d'après le théorème de Sylow, G possède au moins un sous-groupe H de cardinal 2^{d+1}. Alors, $L := K^H$ est de degré r, impair, sur $\mathbb{R}$. Soit $x \in L$ et $Q = \mathrm{Irr}_{\mathbb{R}}(x)$ son polynôme minimal sur $\mathbb{R}$. Alors $\deg Q = [\mathbb{R}(x) : \mathbb{R}]$ divise $[L : \mathbb{R}] = r$ donc est impair.

Or, tout polynôme réel de degré impair a une racine dans $\mathbb{R}$; c'est ici qu'intervient une propriété topologique de $\mathbb{R}$! Donc, Q étant irréductible, il est de degré 1, d'où $x \in \mathbb{R}$. Ceci prouve que $L = \mathbb{R}$ et donc $r = 1$. Par conséquent,

$$[K : \mathbb{C}] = 2^d.$$

Montrons que $d = 0$. Supposons, au contraire, $d \geq 1$. Dans ce cas, $G_1 = \mathrm{Gal}(K/\mathbb{C})$ est un 2-groupe non trivial donc contient un sous-groupe (distingué) H de cardinal 2^{d-1}. Alors, $K' = K^H$ est de degré 2 sur $\mathbb{C}$. Soit $x \in K' \setminus \mathbb{C}$; son polynôme minimal $\mathrm{Irr}_{\mathbb{C}}(x)$ est de degré 2 et irréductible dans $\mathbb{C}[X]$. Mais ceci est une contradiction, puisque dans $\mathbb{C}[X]$, tout polynôme de degré 2 est scindé ! Cette contradiction montre que $d = 0$, d'où $[K : \mathbb{C}] = 1$. Ceci montre que $\mathbb{C}$ est algébriquement clos.

9.4 Discriminant et groupe de Galois d'un polynôme

9.4.1 Signature et groupe alterné A_n

Notation Le groupe $\mathbb{Z}/2\mathbb{Z}$ est aussi noté $\{\pm 1\}$, en notation multiplicative.

Soit $n \geq 2$. On va définir un certain morphisme surjectif $S_n \to \{\pm 1\}$, appelé la **signature**. Ceci peut se faire en utilisant uniquement la structure de groupe de S_n, voir par exemple [ALF, Thm. II.7.3] ou [BR, Chap. 2, §2.9]. Il est commode d'adopter ici l'approche suivante.

Soit k un corps de caractérisque $\neq 2$. Alors $1 \neq -1$ dans k et donc $\{1, -1\}$ est un sous-groupe de $k^\times$, isomorphe à $\mathbb{Z}/2\mathbb{Z}$. On a vu que S_n opère par automorphismes d'algèbre sur $A = k[X_1, \ldots, X_n]$. Considérons le polynôme

$$V_n = \prod_{1 \leq i < j \leq n} (X_i - X_j).$$

Soit $\sigma \in S_n$. Alors σ permute entre eux, au signe près, les facteurs $X_i - X_j$. Plus précisément, on a

$$\sigma(X_i - X_j) = \begin{cases} X_{\sigma(i)} - X_{\sigma(j)}, & \text{si } \sigma(i) < \sigma(j); \\ -(X_{\sigma(j)} - X_{\sigma(i)}), & \text{si } \sigma(i) > \sigma(j). \end{cases}$$

On en déduit que

$$\sigma(V_n) = (-1)^{\ell(\sigma)} V_n,$$

où

$$\ell(\sigma) = |\{i < j \mid \sigma(i) > \sigma(j)\}|$$

est appelé le nombre d'inversions de σ. On pose aussi

$$\varepsilon(\sigma) = (-1)^{\ell(\sigma)};$$

on l'appelle la signature de σ. Observons que $\ell(\mathrm{id}) = 0$. D'autre part, soit τ_{12} la permutation qui échange 1 et 2 et vérifie $\tau_{12}(i) = i$ pour tout $i > 2$. Alors $\ell(\tau_{12}) = 1$ et donc $\varepsilon(\tau_{12}) = -1$. Par conséquent, l'application $\varepsilon : S_n \to \{\pm 1\}$ est surjective.

Proposition 9.4.1 *Pour $n \geq 2$, la signature est un morphisme de groupes $S_n \to \{\pm 1\}$.*

Démonstration. Soient $\sigma, \tau \in S_n$. Alors $(\sigma\tau)(V_n)$ est égal, d'une part, à $\varepsilon(\sigma\tau)V_n$ et, d'autre part, à

$$\sigma(\tau(V_n)) = \sigma(\varepsilon(\tau)V_n) = \varepsilon(\tau)\sigma(V_n) = \varepsilon(\tau)\varepsilon(\sigma)V_n.$$

Il en résulte $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$. Ceci prouve le lemme. □

Définition 9.4.1 *1) On dit qu'une permutation $\sigma \in S_n$ est paire, resp. impaire, si $\varepsilon(\sigma) = 1$, resp. -1.*

2) $\ker \varepsilon$ *est appelé groupe alterné d'ordre n, et noté A_n. Il est formé des permutations paires, et est de cardinal $n!/2$.*

Exemple 9.4.1 Pour $n = 2$, $S_2 \cong \{\pm 1\}$ et $A_2 = \{1\}$. Pour $n = 3$, S_3 est un groupe non-commutatif d'ordre 6, et A_3 est isomorphe à $\mathbb{Z}/3\mathbb{Z}$ et formé des permutations $1 = \mathrm{id}$ et c, $c^2 = c^{-1}$, où

$$c = \begin{pmatrix} 123 \\ 231 \end{pmatrix}, \qquad c^2 = c^{-1} = \begin{pmatrix} 123 \\ 312 \end{pmatrix}.$$

9.4.2 Discriminant d'un polynôme

Soit A un anneau commutatif. On rappelle que l'opérateur de dérivation $D : A[X] \to A[X]$ est l'application A-linéaire définie par $D(1) = 0$ et $D(X^n) = nX^{n-1}$, pour tout $n \geq 1$.

Lemme 9.4.2 *Soient $P_1, \ldots, P_r \in A[X]$. On a*

$$D(P_1 \cdots P_r) = \sum_{i=1}^{r} P_1 \cdots D(P_i) \cdots P_r.$$

Démonstration. Par récurrence sur r. On a déjà vu le cas $r = 2$ (Lemme 7.6.1). Supposons $r \geq 3$ et le résultat établi pour $r - 1$. D'après le cas $r = 2$, l'on a

$$D(P_1 \cdots P_r) = D(P_1)P_2 \cdots P_r + P_1 D(P_2 \cdots P_r),$$

et le résultat découle alors de l'hypothèse de récurrence. □

Théorème 9.4.3 (Discriminant du polynôme général de degré n) *Soient $X_1, \ldots, X_n$ des indéterminées, et $\sigma_1, \ldots, \sigma_n$ les polynômes symétriques élémentaires en $X_1, \ldots, X_n$. Posons $a_i = (-1)^i \sigma_i$. Soient $T_1, \ldots, T_n$ d'autres indéterminées. Il existe un unique polynôme $\Delta_n \in \mathbb{Z}[T_1, \ldots, T_n]$ tel que*

$$\Delta_n(a_1, \ldots, a_n) = \prod_{1 \leq i < j \leq n} (X_i - X_j)^2. \tag{1}$$

Ce polynôme Δ_n est appelé le discriminant du polynôme

$$P = X^n + a_1 X^{n-1} + \cdots + a_n,$$

et est aussi noté disc_P. *De plus, on a*

$$\prod_{i=1}^{n} P'(X_i) = (-1)^{\frac{n(n-1)}{2}} \Delta_n(a_1, \ldots, a_n). \tag{2}$$

Démonstration. Posons $\Pi = \prod_{1 \leq i < j \leq n} (X_i - X_j)^2$; c'est un élément de $\mathbb{Z}[X_1, \ldots, X_n]^{S_n}$. Donc, d'après le théorème fondamental des polynômes symétriques 9.1.3, il existe un unique polynôme $\Delta_n^- \in \mathbb{Z}[T_1, \ldots, T_n]$, tel que

$$\Delta_n^-(\sigma_1, \ldots, \sigma_n) = \Pi.$$

Soit ϕ l'automorphisme de $\mathbb{Z}[T_1, \ldots, T_n]$ défini par $\phi(T_i) = (-1)^i T_i$ pour tout i, et soit $\Delta_n = \phi(\Delta_n^-)$. Alors, Δ_n est l'unique élément de $\mathbb{Z}[T_1, \ldots, T_n]$ vérifiant

$$\Delta_n(a_1, \ldots, a_n) = \Delta_n^-(\sigma_1, \ldots, \sigma_n) = \Pi.$$

Ceci prouve la 1ère assertion.

De plus, on a

$$P = X^n + \sum_{i=1}^{n} (-1)^i \sigma_i X^{n-i} = \prod_{i=1}^{n} (X - X_i).$$

D'après le lemme précédent, on a

$$P' = \sum_{i=1}^{n} \prod_{j \neq i} (X - X_j).$$

Donc, pour tout $i = 1, \ldots, n$, on a $P'(X_i) = \prod_{j \neq i} (X_i - X_j)$. Par conséquent, l'on a

$$\prod_{i=1}^{n} P'(X_i) = \prod_{i \neq j} (X_i - X_j) = (-1)^{\frac{n(n-1)}{2}} \Pi.$$

Ceci prouve le théorème. □

Corollaire 9.4.4 (Discriminant d'un polynome $P \in k[X]$)
Soient k un corps et $P = X^n + \sum_{i=1}^{n} a_i X^{n-i}$ un polynôme unitaire de degré n à coefficients dans k. Soit L une extension de k dans laquelle P est scindé et soient $x_1, \ldots, x_n$ les racines de P dans L. Alors

$$\prod_{1 \leq i < j \leq n} (x_i - x_j)^2 = \Delta_n(a_1, \ldots, a_n).$$

En particulier, P a une racine multiple ssi $\Delta_n(a_1, \ldots, a_n) = 0$.

Démonstration. Plaçons-nous dans l'anneau $R = \mathbb{Z}[X_1, \ldots, X_n]$ et posons $V_n = \prod_{1 \leq i < j \leq n}(X_i - X_j)$ et $A_i = (-1)^i \sigma_i$ pour $i = 1, \ldots, n$. D'après le théorème précédent, on a dans R l'égalité

$$(*) \qquad V_n^2 = \Delta_n(A_1, \ldots, A_n).$$

Soit ϕ l'unique morphisme d'anneaux de R dans L, défini par $\phi(X_i) = x_i$. Pour $r = 1, \ldots, n$, on a

$$\phi(A_r) = (-1)^r \sum_{i_1 < \cdots < i_r} \phi(X_{i_1} \cdots X_{i_r}) = (-1)^r \sigma_r(x_1, \ldots, x_n) = a_r.$$

Par conséquent, appliquant ϕ à l'égalité $(*)$, on obtient

$$\prod_{1 \leq i < j \leq n} (x_i - x_j)^2 = \Delta_n(a_1, \ldots, a_n).$$

La dernière assertion est alors claire. Le corollaire est démontré. □

Proposition 9.4.5 (Discriminant d'un trinôme $X^n + pX + q$)
Soient k un corps et $p, q \in k$. Le discriminant du trinôme $P = X^n + pX + q$, noté disc_P*, égale*

$$(-1)^{n(n-1)/2} \left((1-n)^{n-1} p^n + n^n q^{n-1} \right).$$

En particulier, pour

$$\begin{aligned} P &= X^2 + aX + b, \qquad & \mathrm{disc}_P &= a^2 - 4b; \\ P &= X^3 + pX + q, \qquad & \mathrm{disc}_P &= -4p^3 - 27q^2. \end{aligned}$$

Démonstration. Soit L une extension de k dans laquelle P est scindé et soient $x_1, \ldots, x_n$ les racines de P dans L. D'après l'égalité (2) du théorème 9.4.3, l'égalité à démontrer est équivalente à la suivante :

$$\prod_{i=1}^{n} P'(x_i) = (1-n)^{n-1} p^n + n^n q^{n-1}.$$

Or, $P'(X) = nX^{n-1} + p$. Supposons d'abord $q \neq 0$. Alors, pour $i = 1, \ldots, n$, l'on a $x_i \neq 0$ et

$$(1) \qquad x_i^{n-1} = -p - \frac{q}{x_i}.$$

On en déduit que $\prod_{i=1}^{n} P'(x_i)$ égale

$$(*) \quad (1-n)^n p^n + \frac{(-n)^n q^n}{x_1 \cdots x_n} + \sum_{r=1}^{n-1} (1-n)^r p^r (-nq)^{n-r} \sigma_{n-r}(x_1^{-1}, \ldots, x_n^{-1}).$$

Or, $x_1 \cdots x_n = (-1)^n q$ et, d'autre part, on voit facilement que

$$\sigma_{n-r}(x_1^{-1}, \ldots, x_n^{-1}) = \frac{\sigma_r(x_1, \ldots, x_n)}{x_1 \cdots x_n} = \begin{cases} 0 & \text{si } 1 \leq r \leq n-2; \\ \dfrac{-p}{q} & \text{si } r = n-1. \end{cases}$$

Par conséquent, on déduit de $(*)$ que $\prod_{i=1}^{n} P'(x_i)$ égale

$$(**) \qquad n^n q^{n-1} + (1-n)^{n-1} p^n (1-n+n) = n^n q^{n-1} + (1-n)^{n-1} p^n.$$

Ceci prouve le résultat voulu, lorsque $q \neq 0$. Lorsque $q = 0$, l'argument est analogue : $x_n = 0$ est racine simple, $P'(0) = p$, et pour les autres racines $x_1, \ldots, x_{n-1}$ l'on a $P'(x_i) = (1-n)p$. On obtient ainsi que $\prod_{i=1}^{n} P'(x_i) = (1-n)^{n-1} p^n$ lorsque $q = 0$. Ceci démontre la proposition.

En particulier, pour $n = 2$ ou 3, on obtient que le discriminant du trinôme $X^n + pX + q$ vaut $p^2 - 4q$, resp. $-4p^3 - 27q^2$. □

9.4.3 Groupe de Galois d'un polynôme

Lemme 9.4.6 *Soit G un groupe fini opérant sur un ensemble X et soit $x \in X$. Notons $\mathcal{O}(x)$ l'orbite de x et G_x son stabilisateur. L'application $\phi_x : G \to \mathcal{O}(x)$ induit une bijection $G/G_x \xrightarrow{\sim} \mathcal{O}(x)$ et donc l'on a*

$$|G| = |G_x| \cdot |\mathcal{O}(x)|.$$

Démonstration. Posons $H = G_x$. Pour tout $g \in G$, $h \in H$, on a $\phi_x(gh) = gx = \phi_x(g)$. Par conséquent, ϕ_x induit une application $\psi_x : G/H \to \mathcal{O}(x)$, définie par $\psi_x(gH) = gx$. Cette application est clairement surjective. Reste à voir qu'elle est injective. Pour cela, il faut voir que si $\psi_x(gH) = \psi_x(g'H)$ alors $gH = g'H$. Mais ceci est clair, car si $gx = g'x$ alors $x = g^{-1}g'x$ et donc $g^{-1}g' \in H$, d'où $g' \in gH$ et $g'H = gH$. Ceci prouve la première assertion, et la seconde découle alors du lemme 8.4.1. □

Définition 9.4.2 *Soit G un groupe opérant sur un ensemble X. On dit que l'action est transitive si les éléments de X forment une seule orbite pour l'action de G.*

Théorème 9.4.7 ($\mathrm{Gal}(P/k)$ **est un sous-groupe de** S_n**)**
Soit k un corps et $P \in k[X]$ un polynôme séparable de degré n. Alors :

1) $\mathrm{Gal}(P/k)$ *est un sous-groupe de S_n, donc son ordre divise $n!$.*

2) *Si P est irréductible,* $\mathrm{Gal}(P/k)$ *agit transitivement sur les n racines de P et donc son ordre est divisible par n.*

3) *Plus généralement, écrivons $P = P_1^{m_1} \cdots P_r^{m_r}$, où les P_i sont irréductibles et deux à deux distincts. Posons $d_i = \deg P_i$ et $Q = P_1 \cdots P_r$. Alors* $\mathrm{Gal}(P/k) = \mathrm{Gal}(Q/k)$ *est un sous-groupe de*

$$S_{d_1} \times \cdots \times S_{d_r}.$$

Démonstration. Soit K un corps de rupture de P sur k. On rappelle que l'extension $k \subseteq K$ est galoisienne, d'après le 2ème théorème fondamental 7.4.4, et son groupe de Galois est noté $\mathrm{Gal}(P/k)$. Soient $x_1, \ldots, x_n$ les racines de P dans K.

1) Soit $g \in \mathrm{Gal}(P/k)$. Comme $g(P) = P$, alors $g(x_1), \ldots, g(x_n)$ sont les racines de P dans K ; par conséquent, g induit une permutation $\sigma_g \in S_n$ telle que $g(x_i) = x_{\sigma_g(i)}$ pour tout $i = 1, \ldots, n$. On voit facilement que l'application $g \mapsto \sigma_g$ est un morphisme de groupes. De plus, ce morphisme est injectif puisque les x_i engendrent K sur k. Ceci prouve le 1er point.

2) Si P est irréductible, ses racines $x_1, \ldots, x_n$ sont deux à deux distinctes et forment l'orbite $\mathcal{O}(x_1)$ de x_1 sous $G := \mathrm{Gal}(P/k)$, d'après le corollaire 7.4.5. D'après le lemme précédent, posant

$$H = \mathrm{Stab}_G(x_1) = \{g \in G \mid g(x_1) = x_1\},$$

l'on a $|G| = |H| \cdot |\mathcal{O}(x_1)| = n|H|$. Ceci prouve 2).

3) Plus généralement, écrivons $P = P_1^{m_1} \cdots P_r^{m_r}$ et, pour $i = 1, \ldots, r$, soient $d_i = \deg P_i$ et $\alpha_{i1}, \ldots, \alpha_{id_i}$ les racines de P_i dans K. Soit $g \in G$. Comme $g(P_i) = P_i$, pour tout i, alors g permute les racines de chaque P_i et donc induit une permutation

$$\sigma_g = (\sigma_{g,1}, \ldots, \sigma_{g,r}) \in S_{d_1} \times \cdots \times S_{d_r}$$

telle que $g(\alpha_{ij}) = \alpha_{i\sigma_{g,i}(j)}$ pour tout i, j. Comme précédemment, l'application $g \mapsto \sigma_g$ est un morphisme de groupes, et est injective puisque les α_{ij} engendrent K sur k. Ceci prouve le point 3). □

Remarque 9.4.1 1) Observons aussi que, pour i fixé, les racines α_{ij} sont deux à deux distinctes, puisque P_i est séparable par hypothèse. Par conséquent, $|\mathrm{Gal}(P/k)|$ est divisible par chaque d_i et donc par leur ppcm.

2) De plus, le polynôme minimal de α_{ij} est P_i. Comme $P_i \neq P_{i'}$ pour $i \neq i'$, ceci entraîne que les α_{ij} sont deux à deux distincts.

9.4.4 L'extension intermédiaire associée au discriminant

Soient k un corps de caractéristique $\neq 2$, $P \in k[X]$ un polynôme séparable de degré n, K un corps de décomposition de P sur k, et $G = \mathrm{Gal}(K/k)$. Choisissons une énumération $x_1, \ldots, x_n$ des racines de P dans K et soit ϕ le plongement de G dans S_n défini dans le théorème 9.4.7. Posons

$$d = \prod_{1 \leq i < j \leq n} (x_i - x_j).$$

On a vu que

$$(*) \qquad g(d) = \varepsilon(\phi(g))d, \qquad \forall g \in G.$$

On notera $\varepsilon(g)$ au lieu de $\varepsilon(\phi(g))$. (On peut montrer que $\varepsilon(g)$ ne dépend que de g, et pas de la numérotation $x_1, \ldots, x_n$.)

Théorème 9.4.8 (L'extension intermédiaire $k \subseteq k[d] \subseteq K$)
On suppose $\mathrm{car}(k) \neq 2$. *Soient P, K, ϕ et d comme plus haut. On a : $d \in k \Leftrightarrow \phi(G) \subseteq A_n$. Lorsque $d \notin k$, l'extension $k \subset k[d]$, resp. $k[d] \subseteq K$, est galoisienne, de groupe $\{\pm 1\}$, resp. $\phi(G) \cap A_n$. De plus, $\phi(G) \cap A_n$ est de cardinal $|G|/2$.*

Démonstration. Supposons $\phi(G) \subseteq A_n$. Alors $(*)$ montre que d est invariant par G, donc appartient à k. (Cet argument s'applique également si $\mathrm{car}(k) = 2$.)

Réciproquement, supposons $\phi(G) \not\subseteq A_n = \ker \varepsilon$, et soit $g \in G$ tel que $\phi(g) \notin A_n$. Alors $g(d) = -d$ est différent de d donc $d \notin k$. Posons $\Delta = d^2$. D'après $(*)$, Δ est invariant par G donc appartient à k. (Plus précisément, d'après le corollaire 9.4.4, Δ est le discriminant de P). Le polynôme $X^2 - \Delta$ est séparable sur k, car il a deux racines distinctes d et $-d$. Comme $k[d]$ est le corps de décomposition sur k de $X^2 - \Delta$, on obtient que l'extension $k \subset k[d]$ est galoisienne, de degré 2, et donc de groupe $\{\pm 1\}$.

D'autre part, soit H le fixateur de $k[d]$ dans G. D'après le théorème principal de la théorie de Galois 7.5.7, l'extension $k[d] \subseteq K$ est galoisienne, de groupe H. Or, comme $k[d]$ est engendré sur k par d, l'on a

$$H = \{g \in G \mid g(d) = d\}.$$

Alors, comme $\mathrm{car}(k) \neq 2$, on déduit de $(*)$ que $H = \{g \in G \mid \phi(G) \in A_n\}$, et donc ϕ induit un isomorphisme de H sur $\phi(G) \cap A_n$. Enfin, on obtient que $|H| = |G|/2$, par exemple car $\varepsilon \circ \phi$ induit un isomorphisme $G/H \cong \{\pm 1\}$.

Ou bien, en utilisant le théorème d'Artin et la multiplicativité des degrés, on peut dire que

$$|H| = [K : k(d)] = \frac{[K : k]}{2} = \frac{|G|}{2}.$$

Ceci prouve le théorème. □

Remarque 9.4.2 Dans le théorème précédent, l'hypothèse $\mathrm{car}(k) \neq 2$ est nécessaire. En effet, soit $k = \mathbb{F}_2$. Le polynôme $P = X^2 + X + 1$ a deux racines distinctes dans $\mathbb{F}_4$, car son dérivé est $P' = 1$. Par conséquent,

$$\mathrm{Gal}(P/k) = \mathrm{Gal}(\mathbb{F}_4/\mathbb{F}_2) = \{\pm 1\} = S_2.$$

Pourtant, l'on a $\Delta = 1$ et donc d égale 1 et appartient à k.

9.5 L'équation de degré 3

Soit $Y^3 + aY^2 + bY + c$ un polynôme unitaire de degré 3, à coefficients dans un sous-corps de $\mathbb{C}$. En faisant le changement de variable $X = Y + a/3$, on se ramène à l'équation

$$0 = X^3 + pX + q, \tag{1}$$

où $p = b - a^2/3$ et $q = 2a^3/27 - ba/3 + c$. Si $p = 0$, les racines de (1) sont les racines cubiques de $-q$; on supposera donc dans la suite $p \neq 0$.

L'équation (1) a été résolue au XVIe siècle, voir par exemple [Ti] pour une discussion historique. En language moderne, on peut présenter cette solution comme suit. Cherchons X sous la forme $X = y + z$, où y, z sont deux indéterminées auxiliaires. Alors, (1) équivaut à

$$y^3 + z^3 + (3yz + p)(y + z) + q = 0. \tag{2}$$

Par conséquent, si l'on pose $z = -p/3y$, on obtient l'équation

$$y^3 - \frac{p^3}{27y^3} + q = 0,$$

d'où, en multipliant par y^3, l'équation

$$y^6 + qy^3 - \left(\frac{p}{3}\right)^3 = 0. \tag{3}$$

Par conséquent, y^3 est racine de l'équation du second degré

$$T^2 + qT - \left(\frac{p}{3}\right)^3 = 0. \tag{4}$$

De façon plus symétrique, on peut dire que si l'on impose $yz = -p/3$, alors y^3 et z^3 sont solutions de

$$y^3 z^3 = -(p/3)^3 \quad \text{et} \quad y^3 + z^3 = -q,$$

donc sont les racines de l'équation du second degré (4). Quitte à permuter y et z, on peut donc écrire

$$\left\{ \begin{array}{rcl} y^3 & = & -\frac{q}{2} + \frac{1}{2}\sqrt{q^2 + 4\left(\frac{p}{3}\right)^3} = -\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3} = A, \\ z^3 & = & -\frac{q}{2} \pm \frac{1}{2}\sqrt{q^2 + 4\left(\frac{p}{3}\right)^3} = -\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3} = B. \end{array} \right.$$

Observons que $AB \neq 0$, puisqu'on a supposé $p \neq 0$. Soit α l'une des racines cubiques dans $\mathbb{C}$ de A ; les deux autres sont $j\alpha$ et $j^2\alpha$, où $j = \exp(2i\pi/3)$. Soit β la racine cubique de B déterminée par la condition $\alpha\beta = -p/3$, c.-à-d., $\beta = -p/3\alpha$. Alors, les racines de l'équation (1) sont

$$\left\{ \begin{array}{rcl} x_1 & = & \alpha + \beta, \\ x_2 & = & j\alpha + j^2\beta, \\ x_3 & = & j^2\alpha + j\beta. \end{array} \right. \tag{5}$$

Remarque 9.5.1 Posons $P = X^3 + pX + q$. Soit $K = \mathbb{Q}(p, q)$ le sous-corps de $\mathbb{C}$ engendré par les coefficients de P et soit L le sous-corps de $\mathbb{C}$ engendré par les racines x_1, x_2, x_3 de P dans $\mathbb{C}$ (il contient K puisque $p = x_1x_2 + x_1x_3 + x_2x_3$ et $q = -x_1x_2x_3$). Posons

$$\Delta = -27q^2 - 4p^3 = -3(3 \cdot 2)^2\Delta';$$

c'est le discriminant de P. On a vu précédemment que $\Delta = [(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)]^2$; par conséquent l'élément

$$(x_1 - x_2)(x_1 - x_3)(x_2 - x_3) \in L$$

est une racine carrée de Δ.

1) Les formules (5) montrent que les racines de P s'écrivent comme somme de racines cubiques de $-q/2 \pm \sqrt{\Delta'}$. Mais attention, en général ces racines cubiques n'appartiennent pas à L. Toutefois, on verra plus loin que

ces racines cubiques sont dans L si le sous-corps $K[\sqrt{\Delta}]$ contient $i\sqrt{3}$ ou, de façon équivalente, $j = (-1 + i\sqrt{3})/2$.

2) Supposons P irréductible sur K. Alors $G := \mathrm{Gal}(L/K)$ est un sous-groupe de S_3 d'ordre divisible par 3. Si $d = \sqrt{\Delta}$ appartient à K alors, d'après le théorème 9.4.8, G est contenu dans $A_3 \cong \mathbb{Z}/3\mathbb{Z}$, et donc $G = A_3$. D'autre part, si $d \notin K$, alors $|G| = [L : K]$ est divisible par $[K[d] : K] = 2$, d'où $|G| = 6$ et $G = S_3$. Par conséquent, on a, pour P irréductible de degré 3,

$$\begin{cases} \mathrm{Gal}(P/K) \cong A_3 \cong \mathbb{Z}/3\mathbb{Z} & \text{si } \sqrt{\Delta} \in K; \\ \mathrm{Gal}(P/K) \cong S_3 & \text{si } \sqrt{\Delta} \notin K. \end{cases}$$

Table des matières

Bibliographie

[] Voici une bibliographie provisoire (elle aussi en évolution au fil du texte).

[ALF] J.-M. Arnaudiès, J. Lelong-Ferrand, Cours de mathématiques, Tome 1, Algèbre (3ème édition), Dunod, 1978.

[AM] M. Atiyah, I. G. Macdonald, Commutative algebra, Addison-Wesley, 1969.

[Art] E. Artin, Galois Theory, nouvelle édition, Dover, 1998.

[Bla] A. Blanchard, Les corps non commutatifs, P.U.F., 1972.

[BAlg] N. Bourbaki, Algèbre, Chapitres 4 à 7, Masson, 1981.

[BR] A. Bouvier, D. Richard, Groupes (2ème édition revue et corrigée), Hermann, 1979.

[BM] J. Briançon, Ph. Maisonobe, Éléments d'algèbre commutative (niveau M1), Ellipses, 2004.

[Die] J. Dieudonné, Cours de géométrie algébrique, tome 2, P.U.F., 1974.

[Dou] A. Douady, R. Douady, Algèbre et théories galoisiennes (2 tomes), Cedic Fernand Nathan, 1977.

[Esc] J.-P. Escofier, Théorie de Galois, Dunod, 2000.

[Ja1] N. Jacobson, Basic algebra I, W. H. Freeman & Co., 1974.

[Ja2] N. Jacobson, Basic algebra II, W. H. Freeman & Co., 1980.

[Kri] J.-L. Krivine, Théorie des ensembles, Cassini, 1998.

[Ku] E. Kunz, Introduction to commutative algebra and algebraic geometry, Birkhäuser, 1985.

[La] S. Lang, Algebra, Addison-Wesley, 1965. Traduction française de la 3ème édition : *Algèbre*, Dunod, 2004.

[Laf] J.-P. Lafon, Les formalismes fondamentaux de l'algèbre commutative, Hermann, 1974.

[Pe1] D. Perrin, Cours d'algèbre, E.N.S.J.F. 1981, et 3ème édition, Ellipses, 1996.

[Pe2] D. Perrin, Géométrie algébrique - Une introduction, Inter Éditions/-CNRS Éditions, 1995.

[Sa] P. Samuel, Théorie algébrique des nombres, Hermann, 1967.

[Se] J.-P. Serre, Représentations linéaires des groupes finis (3ème édition corrigée), Hermann, 1978.

[SD] H.P.F. Swinnerton-Dyer, A brief guide to algebraic number theory, C.U.P., 2001.

[Ti] J.-P. Tignol, Galois' Theory of algebraic equations, World Scientific, 2001.