LES POLYNÔMES SYMÉTRIQUES

1. Les polynômes symétriques élémentaires

Fixons quelques notations et terminologies pour les polynômes à plusieurs variables. On fixe un corps K. Un *polynôme* en $X_1, X_2, \ldots, X_n$ avec coefficients dans K est une somme

$$P(X) = \sum_{(i_1,\ldots,i_n)\in\mathbb{N}^n} a_{i_1,\ldots,i_n} X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n} \tag{1}$$

avec les $a_{i_1,\ldots,i_n} \in K$ et tous sauf un nombre fini des $a_{i_1,\ldots,i_n}$ égaux à 0. L'ensemble de polynômes en $X_1, \ldots, X_n$ avec coefficients dans K forme un anneau commutatif $K[X_1, \ldots, X_n]$.

Les produits $a_{i_1,\ldots,i_n} X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}$ avec $a_{i_1,\ldots,i_n} \in K^\times$ sont des *termes*. Les $X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}$ sont des *monômes*. Donc un polynôme est une somme d'un nombre fini de termes et une combinaison linéaire d'un nombre fini de monômes.

Définition. Le *degré* d'un terme ou monôme $a_{i_1,\ldots,i_n} X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}$ est $i_1 + i_2 + \cdots + i_n$. Le *degré* d'un polynôme non nul est le degré maximal de ses termes.

Un polynôme non nul est *homogène* de degré d si tous ses termes sont de degré d. Le polynôme 0 est homogène de tout degré.

Par exemple $X_1 X_3^3 + X_2^2 X_4^2 + X_5^4$ est homogène de degré 4.

Les polynômes homogènes d'un degré donné forment un espace vectoriel vectoriel. Les polynômes homogènes de degré 1 sont les *formes linéaires*

$$\sum_{i=1}^{n} a_i X_i = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n.$$

Les polynômes homogènes de degré 2 sont les *formes quadratiques*

$$\sum_{1\le i\le j\le n} a_{ij} X_i X_j = a_{11} X_1^2 + a_{12} X_1 X_2 + \cdots + a_{nn} X_n^2.$$

Parfois on parle de formes cubiques, quartiques, etc.

Strictement dit le polynôme 0 n'a pas de degré parce qu'il n'a pas de termes, mais parfois il est convenable de poser $\deg(0) = -1$ ou $\deg(0) = -\infty$.

Dans ce cours nous étudierons principalement les polynômes symétriques.

Définition. Un polynôme $P \in K[X_1, X_2, \ldots, X_n]$ en n variables est *symétrique* si pour toute permutation $\rho \in S_n$ on a

$$P(X_1, X_2, \ldots, X_n) = P(X_{\rho(1)}, X_{\rho(2)}, \ldots, X_{\rho(n)}).$$

Un polynôme $P(X_1, \ldots, X_n) = \sum a_{i_1,\ldots,i_n} X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}$ est symétrique ssi pour toute permutation $\rho \in S_n$ et toute multi-indice $(i_1, \ldots, i_n) \in \mathbb{N}^n$ on a

$$a_{i_1,i_2,\ldots,i_n} = a_{i_{\rho(1)},i_{\rho(2)},\ldots,i_{\rho(n)}}$$

Par exemple $X_1 + X_2$ et $X_1^2 + 4X_1X_2 + X_2^2$ sont des polynômes symétriques en deux variables, et $X_1^2X_2 + X_1^2X_3 + X_1X_2^2 + X_1X_3^2 + X_2^2X_3 + X_2X_3^2$ est un polynôme symétrique en trois variables.

Définition. Pour $1 \leq k \leq n$ le *k-ième polynôme symétrique élémentaire* en n variables est

$$\sigma_i(X_1, \dots, X_n) = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} X_{i_1} X_{i_2} \cdots X_{i_k} \tag{2}$$

Ainsi σ_k est la somme de tous les produits de k variables distincts. Il est homogène de degré k. Il est la somme de $\binom{n}{k} = C_n^k$ monômes. Par exemple les polynômes symétriques élémentaires en trois variables sont

$$\sigma_1 = X_1 + X_2 + X_3, \qquad \sigma_2 = X_1X_2 + X_1X_3 + X_2X_3, \qquad \sigma_3 = X_1X_2X_3. \tag{3}$$

Théorème 1.1. *Soit $r_1, \dots, r_n \in K$. Pour $k = 1, \dots, n$ soit $\sigma_k = \sigma_k(r_1, \dots, r_n)$ le k-ième polynôme symétrique élémentaire en $r_1, \dots, r_n$. Alors on a*

$$\prod_{i=1}^{n} (T - r_i) = T^n - \sigma_1 T^{n-1} + \sigma_2 T^{n-2} - \cdots + (-1)^n \sigma_n.$$

Par exemple, les polynômes symétriques élémentaires en 1, 2 et 3 sont

$$\sigma_1 = 1 + 2 + 3 = 6, \qquad \sigma_2 = 1 \cdot 2 + 1 \cdot 3 + 2 \cdot 3 = 11, \qquad \sigma_3 = 1 \cdot 2 \cdot 3 = 6,$$

et on a

$$(T-1)(T-2)(T-3) = T^3 - 6T^2 + 11T - 6.$$

Les polynômes symétriques élémentaires en u, $-u$, v et $-v$ sont

$$\begin{aligned}
\sigma_1 &= u + (-u) + v + (-v) = 0, \\
\sigma_2 &= u(-u) + uv + u(-v) + (-u)v + (-u)(-v) + v(-v) = -u^2 - v^2, \\
\sigma_3 &= u(-u)v + u(-u)(-v) + uv(-v) + (-u)v(-v) = 0, \\
\sigma_4 &= u(-u)v(-v) = u^2v^2,
\end{aligned}$$

et on a bien

$$(T-u)(T+u)(T-v)(T+v) = (T^2 - u^2)(T^2 - v^2) = T^4 - (u^2 + v^2)T^2 + u^2v^2.$$

Démonstration du théorème 1.1. Quand on développe un produit de n facteurs, et chaque facteur est la somme de 2 termes, on trouve une somme de 2^n termes indexés par les parties $I \subset \{1, 2, \dots, n\}$

$$\prod_{i=1}^{n} (A_i + B_i) = \sum_{I \subset \{1,2,\dots,n\}} \prod_{i \notin I} A_i \prod_{i \in I} B_i.$$

(Cette formule se démontre par récurrence sur n.) Donc nous avons

$$\prod_{i=1}^{n} (T - r_i) = \sum_{I \subset \{1,2,\dots,n\}} \prod_{i \notin I} T \prod_{i \in I} (-r_i)$$

Dans cette somme le terme correspondant à la partie $I = \{i_1, \dots, i_k\}$ de cardinal k est égal à $(-1)^k r_{i_1} r_{i_2} \cdots r_{i_k} T^{n-k}$. Donc en regroupant les termes selon les cardinaux $k = |I|$ on a

$$\prod_{i=1}^{n} (T - r_i) = \sum_{k=0}^{n} (-1)^k \Big(\sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} r_{i_1} r_{i_2} \cdots r_{i_k} \Big) T^{n-k} = \sum_{k=0}^{n} (-1)^k \sigma_k T^{n-k}$$

avec $\sigma_0 = 1$ et $\sigma_k = \sigma_k(r_1, \dots, r_n)$ pour $1 \leq k \leq n$. □

Le reste du paragraphe sera dédié à la démonstration du théorème suivant.

Théorème 1.2. *Tout polynôme symétrique dans $K[X_1, X_2, \dots, X_n]$ s'écrit d'une façon unique comme une expression polynomiale en les polynômes symétriques élémentaires $\sigma_1, \sigma_2, \dots, \sigma_n$.*

C'est-à-dire pour tout polynôme symétrique $P \in K[X_1, X_2, \dots, X_n]$ il existe un unique polynôme Q en n variables tel qu'on ait

$$P(X_1, \dots, X_n) = Q(\sigma_1, \dots, \sigma_n).$$

Par exemple dans $K[X_1, X_2, X_3]$ on a

$$X_1^2 + X_2^2 + X_3^2 = \sigma_1^2 - 2\sigma_2,$$
$$X_1^2X_2 + X_1^2X_3 + X_1X_2^2 + X_1X_3^2 + X_2^2X_3 + X_2X_3^2 = \sigma_1\sigma_2 - 3\sigma_3.$$

Ces formules se vérifient en substituant les formules (3) pour σ_1, σ_2 et σ_3 dans les membres de droite et en développant le résultat.

Nous démontrerons la partie *Existence* du théorème 1.2 en donnant un algorithme qui pour chaque polynôme symétrique $P(X_1, \dots, X_n)$ trouve le $Q(\sigma_1, \dots, \sigma_n)$ qui lui est égal. Mais avant cela nous devons développer plusieurs notions.

En travaillant avec un polynôme d'une variable $a_dX^d + \cdots + a_1X + a_0$, on regarde souvent son terme dominant, qui est a_dX^d (si on a $a_d \neq 0$). Pour faire quelque chose similaire avec les polynômes de plusieurs variables, il faut ordonner tous les monômes.

Définition. Un monôme (ou un terme) est avant un autre dans l'*ordre lexicographique*, noté

$$X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n} \succ X_1^{s_1}X_2^{s_2}\cdots X_n^{s_n}$$

si la première fois qu'on a $r_i \neq s_i$ on a $r_i > s_i$. C'est-à-dire, s'il existe un m avec $r_i = s_i$ pour $i < m$ et avec $r_m > s_m$.

Cet ordre a plusieurs propriétés :

(i) Transitivité : Si on a

$$X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n} \succ X_1^{s_1}X_2^{s_2}\cdots X_n^{s_n} \qquad \text{et} \qquad X_1^{s_1}X_2^{s_2}\cdots X_n^{s_n} \succ X_1^{t_1}X_2^{t_2}\cdots X_n^{t_n},$$

alors on a $X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n} \succ X_1^{t_1}X_2^{t_2}\cdots X_n^{t_n}$.

(ii) Trichotomie : Pour chaque couple de monômes exactement un des trois énoncés suivants est vrai :

$$X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n} \succ X_1^{s_1}X_2^{s_2}\cdots X_n^{s_n},$$
$$X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n} = X_1^{s_1}X_2^{s_2}\cdots X_n^{s_n},$$
$$X_1^{s_1}X_2^{s_2}\cdots X_n^{s_n} \succ X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n}.$$

(iii) Compatibilité avec la multiplication : Si on a $X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n} \succ X_1^{s_1}X_2^{s_2}\cdots X_n^{s_n}$, alors en multipliant par un monôme $X_1^{t_1}X_2^{t_2}\cdots X_n^{t_n}$, on garde

$$X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n} \cdot X_1^{t_1}X_2^{t_2}\cdots X_n^{t_n} \succ X_1^{s_1}X_2^{s_2}\cdots X_n^{s_n} \cdot X_1^{t_1}X_2^{t_2}\cdots X_n^{t_n} \tag{4}$$

(iv) Pour toute variable on a $X_i \succ 1$.

(v) Compatibilité avec l'ordre des variables : $X_1 \succ X_2 \succ \cdots \succ X_n$.

Les axiomes (i)–(ii) décrivent un *ordre total*. Une relation sur les monômes de $K[X_1, \dots, X_n]$ vérifiant (i)–(iv) est un *ordre monomial*. La condition (v) est considérée plutôt comme une convénience que comme un axiome fondamental. Dans la littérature l'axiome (iv) est parfois remplacé par des axiomes équivalents.

Rien ne change dans la suite si on utilise un autre ordre monomial vérifiant (v) dans la place de l'ordre lexicographique.

Définition. Le *terme initial* d'un polynôme non nul $P \in K[X_1, X_2, \dots, X_n]$ est le terme de P qui est avant tous les autres termes de P dans l'ordre. On le note $\mathrm{in}(P)$.

Par exemple dans $P = -2X_1^2X_3 + 3X_1X_2X_3 + X_2^3$ on a trié les trois termes pour qu'ils apparaissent dans l'ordre lexicographique. Le terme initial est $\mathrm{in}(P) = -2X_1^2X_3$ parce qu'il est avant $3X_1X_2X_3$ et X_2^3 dans l'ordre lexicographique.

Lemme 1.3. *Soit $P, Q \in K[X_1, X_2, \dots, X_n]$ des polynômes non nuls.*
(a) On a $\mathrm{in}(PQ) = \mathrm{in}(P)\,\mathrm{in}(Q)$.
(b) Si on a $\mathrm{in}(P) = \mathrm{in}(Q)$, alors on a $\mathrm{in}(P) \succ \mathrm{in}(P - Q)$ si on a $P - Q \neq 0$.

Preuve. (a) Les termes de PQ sont des sommes de produits de termes de P et de Q. Pour tout autre terme t_P de P et tout autre terme t_Q de Q on a $\mathrm{in}(P) \succ t_P$ et $\mathrm{in}(Q) \succ t_Q$. Donc la formule (4) nous donne

$$\mathrm{in}(P)\,\mathrm{in}(Q) \succ \mathrm{in}(P)t_Q \succ t_Pt_Q, \qquad \mathrm{in}(P)\,\mathrm{in}(Q) \succ t_P\,\mathrm{in}(Q).$$

Donc le produit de termes $\mathrm{in}(P)\,\mathrm{in}(Q)$ est avant tout autre produit de termes de P et de Q. Il est le terme initial $\mathrm{in}(PQ)$.

(b) Si on a $\mathrm{in}(P) = \mathrm{in}(Q)$, alors ces deux termes s'annulent l'un contre l'autre dans $P - Q$, et le terme initial de $P - Q$ (si cette différence est non nulle) provient d'un terme non initial de P ou de Q. Donc $\mathrm{in}(P - Q)$ est après $\mathrm{in}(P)$ dans l'ordre.

(c) Si on a $\mathrm{in}(P) \succ \mathrm{in}(Q_i)$ pour tout i, alors $\mathrm{in}(P)$ est avant les autres termes de P et tous les termes de tous les Q_i. Par conséquence $\mathrm{in}(P)$ est le terme initial $\mathrm{in}(P + \sum_i Q_i)$. □

Lemme 1.4. *Soit $P \in K[X_1, X_2, \dots, X_n]$ un polynôme symétrique non nul avec terme initial $\mathrm{in}(P) = aX_1^{r_1}X_2^{r_2}\cdots X_n^{r_n}$. Alors on a $r_1 \geq r_2 \geq \cdots \geq r_n \geq 0$.*

L'idée est que dans un polynôme symétrique P un terme comme $3X_1X_3^4X_4^2$ ne peut pas être le terme initial parce que s'il apparaît dans P, alors $3X_1^4X_2^2X_3$ y apparaît aussi car P est symétrique, et on a $3X_1^4X_2^2X_3 \succ 3X_1X_3^4X_4^2$. Donc dans le terme initial $\mathrm{in}(P) = aX_1^{r_1}X_2^{r_2}\cdots X_n^{r_n}$ les puissances des variables successives sont triées en ordre décroissant $r_1 \geq r_2 \geq \cdots \geq r_n$.

Lemme 1.5. *Soit $r_1 \geq r_2 \geq \cdots \geq r_n \geq 0$ entiers, et posons $Q = \sigma_1^{r_1-r_2}\sigma_2^{r_2-r_3}\cdots\sigma_{n-1}^{r_{n-1}-r_n}\sigma_n^{r_n}$. Alors on a $\mathrm{in}(Q) = X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n}$.*

En plus, si on a $(s_1, s_2, \dots, s_n) \neq (t_1, \dots, t_n)$ dans $\mathbb{N}^n$, alors les monômes $\mathrm{in}(\sigma_1^{s_1}\sigma_2^{s_2}\cdots\sigma_n^{s_n})$ et $\mathrm{in}(\sigma_1^{t_1}\sigma_2^{t_2}\cdots\sigma_n^{t_n})$ sont distincts.

Preuve. Les termes initiaux des polynômes symétriques élémentaires sont $\mathrm{in}(\sigma_i) = X_1X_2\cdots X_i$ parce que la suite de i fois 1's et $n - i$ fois 0 vérifiant la condition de décroissance du lemme 1.4 est $(1, \dots, 1, 0, \dots, 0)$. Par la proposition 1.3(a) on a

$$\mathrm{in}(Q) = X_1^{r_1-r_2}(X_1X_2)^{r_2-r_3}\cdots(X_1\cdots X_{n-1})^{r_{n-1}-r_n}(X_1\cdots X_{n-1}X_n)^{r_n}.$$

La puissance totale de X_i dans $\mathrm{in}(Q)$ est

$$(r_i - r_{i+1}) + (r_{i+1} - r_{i+2}) + \cdots + (r_{n-1} - r_n) + r_n = r_i.$$

Donc on a bien $\mathrm{in}(Q) = X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n}$.

Par un calcul similaire

$$\mathrm{in}(\sigma_1^{s_1}\sigma_2^{s_2}\cdots\sigma_n^{s_n}) = X_1^{\sum_{i=1}^n s_i}X_2^{\sum_{i=2}^n s_i}\cdots X_{n-1}^{s_{n-1}+s_n}X_n^{s_n}.$$

Si on a $\mathrm{in}(\sigma_1^{s_1}\sigma_2^{s_2}\cdots\sigma_n^{s_n}) = \mathrm{in}(\sigma_1^{t_1}\sigma_2^{t_2}\cdots\sigma_n^{t_n})$, alors on a

$$s_k = \sum_{i=k}^{n} s_i - \sum_{i=k+1}^{n} s_i = \sum_{i=k}^{n} t_i - \sum_{i=k+1}^{n} t_i = t_k$$

pour tout k. On en déduit $(s_1, s_2, \ldots, s_n) = (t_1, t_2, \ldots, t_n)$. □

Exemple. Illustrons comment écrire

$$P_0 = X_1^3 + X_1^2X_2 + X_1^2X_3 + X_1X_2^2 + X_1X_2X_3 + X_1X_3^2 + X_2^3 + X_2^2X_3 + X_2X_3^2 + X_3^3.$$

comme en termes des polynômes symétriques élémentaires σ_1, σ_2 et σ_3. Pour simplifier l'écriture écrivons

$$\begin{aligned} X_1^3 + \cdots &= X_1^3 + X_2^3 + X_3^3, \\ X_1^2X_2 + \cdots &= X_1^2X_2 + X_1^2X_3 + X_1X_2^2 + X_1X_3^2 + X_2^2X_3 + X_2X_3^2 \end{aligned}$$

Notre polynôme est alors

$$P_0 = (X_1^3 + \cdots) + (X_1^2X_2 + \cdots) + X_1X_2X_3.$$

Son terme initial est $\mathrm{in}(P) = X_1^3$. Par le lemme 1.5 on a aussi $\mathrm{in}(\sigma_1^3) = X_1^3$. On pose

$$\begin{aligned} P_1 &= P_0 - \sigma_1^3 \\ &= \big((X_1^3 + \cdots) + (X_1^2X_2 + \cdots) + X_1X_2X_3\big) \\ &\qquad - \big((X_1^3 + \cdots) + 3(X_1^2X_2 + \cdots) + 6X_1X_2X_3\big) \\ &= -2(X_1^2X_2 + \cdots) - 5X_1X_2X_3 \end{aligned}$$

Ce polynôme a $\mathrm{in}(P_1) = -2X_1^2X_2$. Par le lemme on a $\mathrm{in}(2\sigma_1\sigma_2) = 2X_1^2X_2$. On pose

$$\begin{aligned} P_2 &= P_1 + 2\sigma_1\sigma_2 = P_0 - \sigma_1^3 + 2\sigma_1\sigma_2 \\ &= \big(-2(X_1^2X_2 + \cdots) - 5X_1X_2X_3\big) + \big(2(X_1^2X_2 + \cdots) + 6X_1X_2X_3\big) \\ &= X_1X_2X_3 \end{aligned}$$

On a $\mathrm{in}(P_2) = X_1X_2X_3$ et $\mathrm{in}(\sigma_3) = X_1X_2X_3$. Donc on pose

$$\begin{aligned} P_3 &= P_2 - \sigma_3 = P_0 - \sigma_1^3 + 2\sigma_1\sigma_2 - \sigma_3 \\ &= X_1X_2X_3 - X_1X_2X_3 \\ &= 0. \end{aligned}$$

On trouve donc $P_0 - \sigma_1^3 + 2\sigma_1\sigma_2 - \sigma_3 = 0$ et par conséquent $P_0 = \sigma_1^3 - 2\sigma_1\sigma_2 + \sigma_3$.

Algorithme pour écrire un polynôme symétrique $P(X_1, \ldots, X_n)$ sous la forme $Q(\sigma_1, \ldots, \sigma_n)$. L'algorithme calcule une suite de polynômes symétriques P_i et une suite de polynômes Q_i avec

$$P_i = P - Q_i(\sigma_1, \ldots, \sigma_n).$$

On commence en posant

$$P_0 = P, \qquad\qquad Q_0 = 0.$$

Maintenant supposons qu'on a calculé P_i et Q_i. Il y a deux cas :

Si $P_i = 0$, on pose $Q = Q_i$. **STOP**

Si $P_i \neq 0$, on cherche son terme initial. Selon le lemme 1.4 ce terme initial s'écrit sous la forme $\mathrm{in}(P) = aX_1^{r_1}X_2^{r_2}\cdots X_n^{r_n}$ avec $r_1 \geq r_2 \geq \cdots \geq r_n \geq 0$. On pose

$$P_{i+1} = P_i - a\sigma_1^{r_1-r_2}\sigma_2^{r_2-r_3}\cdots\sigma_{n-1}^{r_{n-1}-r_n}\sigma_n^{r_n}$$

et on développe P_{i+1} comme un polynôme en $X_1, \ldots, X_n$ en utilisant les formules pour les σ_i en termes de $X_1, \ldots, X_n$. On pose

$$Q_{i+1} = Q_i + a\sigma_1^{r_1-r_2}\sigma_2^{r_2-r_3}\cdots\sigma_{n-1}^{r_{n-1}-r_n}\sigma_n^{r_n}$$

sans évaluer les σ_j. (Dans les Q_i on traite les σ_j comme s'ils étaient des variables.)

On répète ce boucle calculant les P_i et Q_i successifs jusqu'a ce qu'on trouve un N avec $P_N = 0$ et donc $Q = Q_N$.

L'algorithme se termine pour les raison suivantes. Ecrivons $R = a\sigma_1^{r_1-r_2}\sigma_2^{r_2-r_3}\cdots\sigma_{n-1}^{r_{n-1}-r_n}$. Par le lemme 1.5 on a $\mathrm{in}(R) = aX_1^{r_1}X_2^{r_2}\cdots X_n^{r_n} = \mathrm{in}(P_i)$. Donc par le lemme 1.3(b) on a $\mathrm{in}(P_i) \succ \mathrm{in}(P_i - R) = \mathrm{in}(P_{i+1})$ si $P_{i+1} \neq 0$. Donc on a

$$\mathrm{in}(P_0) \succ \mathrm{in}(P_1) \succ \mathrm{in}(P_2) \succ \cdots$$

tant que les P_i sont non nuls. Donc il n'y a pas de répétitions parmi monômes initiaux des P_i.

Mais d'autre part si $\deg(P) = d$, alors tous les monômes et polynômes qui apparaissent dans le déroulement de l'algorithme sont de degré $\leq d$. Donc les monômes initiaux des P_i appartiennent à un ensemble fini.

Donc la suite des P_i non nuls doit s'arrêter, et on doit tomber sur un N avec $P_N = 0$.

Preuve du théorème 1.2. L'algorithme ci-dessus montre la partie *Existence* du théorème : pour tout polynôme symétrique $P \in K[X_1, \ldots, X_n]$ il existe un polynôme Q en n variables avec $P = Q(\sigma_1, \ldots, \sigma_n)$.

Unicité : L'unicité de Q est équivalent à ce que l'application

$$\begin{aligned} K[T_1, \ldots, T_n] &\longrightarrow K[X_1, \ldots, X_n] \\ Q(T_1, \ldots, T_n) &\longmapsto Q(\sigma_1, \ldots, \sigma_n) \end{aligned}$$

soit injectif. Mais cette application est un morphisme d'anneaux, et un morphisme d'anneaux est injectif ssi son noyau est $\{0\}$.

Donc il suffit de montrer que pour $Q(T_1, \ldots, T_n) \neq 0$ on a aussi $Q(\sigma_1, \ldots, \sigma_n) \neq 0$. Un tel Q s'écrit $Q = \sum_{j=1}^{M} a_j T_1^{s_{j1}} T_2^{s_{j2}} \cdots T_n^{s_{jn}}$. On peut supposer trois hypothèses :

(i) Les monômes apparaissant dans les différents termes de la somme sont distincts, c'est à dire $(s_{j1}, s_{j2}, \ldots, s_{jn}) \neq (s_{k1}, s_{k2}, \ldots, s_{kn})$ pour $j \neq k$.

(ii) Tous les coefficients sont non nuls : $a_j \neq 0$.

(iii) Il y a au moins un terme dans la somme : $M \geq 1$.

Selon le lemme 1.5 l'hypothèse (i) implique que les monômes initiaux $\mathrm{in}(\sigma_1^{s_{j1}}\sigma_2^{s_{j2}}\cdots\sigma_n^{s_{jn}})$ sont distincts. Il y en a au moins 1 par l'hypothèse (iii), donc il y en a un $\mathrm{in}(\sigma_1^{s_{j_0 1}}\sigma_2^{s_{j_0 2}}\cdots\sigma_n^{s_{j_0 n}})$ qui est avant tous les autres. Alors $a_{j_0}\,\mathrm{in}(\sigma_1^{s_{j_0 1}}\sigma_2^{s_{j_0 2}}\cdots\sigma_n^{s_{j_0 n}})$ est le terme initial non nul de $Q(\sigma_1, \ldots, \sigma_n)$ par le lemme 1.3(c). En particulier on a $Q(\sigma_1, \ldots, \sigma_n) \neq 0$. □

Exemple. Soit r_1, r_2 et r_3 les racines dans $\mathbb{C}$ du polynôme $f(x) = x^3 - 6x^2 + 2x + 2$. Quel est le polynôme $g(x)$ avec racines $r_1 + r_2$, $r_1 + r_3$ et $r_2 + r_3$?

Solution. Selon le théorème 1.1 on a

$$\sigma_1 = r_1 + r_2 + r_3 = 6, \qquad \sigma_2 = r_1r_2 + r_1r_3 + r_2r_3 = 2, \qquad \sigma_3 = r_1r_2r_3 = -2.$$

Le polynôme unitaire avec racines r_1+r_2, r_1+r_3 et r_2+r_3 est $g(x)=x^3-a_1x^2+a_2x-a_3$ avec

$$\begin{aligned}
a_1 &= (r_1+r_2)+(r_1+r_3)+(r_2+r_3) = 2r_1+2r_2+2r_3 = 2\sigma_1 = 12,\\
a_2 &= (r_1+r_2)(r_1+r_3)+(r_1+r_2)(r_2+r_3)+(r_1+r_3)(r_2+r_3)\\
&\quad = r_1^2+r_2^2+r_3^2+3r_1r_2+3r_1r_3+3r_2r_3 = \sigma_1^2+\sigma_2 = 6^2+2 = 38,\\
a_3 &= (r_1+r_2)(r_1+r_3)(r_2+r_3) = r_1^2r_2+r_1^2r_3+r_1r_2^2+r_1r_3^3+r_2^2r_3+r_2r_3^2+2r_1r_2r_3\\
&\quad = \sigma_1\sigma_2-\sigma_3 = 2\cdot 6-(-2) = 14.
\end{aligned}$$

Donc le polynôme avec racines r_1+r_2, r_1+r_3 et r_2+r_3 est $g(x)=x^3-12x^2+38x-14$.

2. La dérivée formelle d'un polynôme

Définition. La *dérivée formelle* d'un polynôme

$$P(X)=a_nX^n+a_{n-1}X^{n-1}+\cdots+a_2X^2+a_1X+a_0$$

dans $K[X]$ est

$$P'(X)=na_nX^{n-1}+(n-1)a_{n-1}X^{n-1}+\cdots+2a_2X+a_1.$$

La dérivation de polynômes est ***linéaire*** et satisfait à la règle de Leibniz :

$$\big(P(X)+Q(X)\big)'=P'(X)+Q'(X) \qquad \big(aP(X)\big)'=aP'(X)$$
$$\big(P(X)Q(X)\big)'=P'(X)Q(X)+P(X)Q'(X).$$

Les deux membres de l'équation de Leibniz sont bilinéaires en $P(X)$ et $Q(X)$, donc il suffit de vérifier la règle pour les monômes, et on a bien

$$(X^n\cdot X^m)'=(X^{n+m})'=(n+m)X^{n+m-1}=nX^{n-1}\cdot X^m+X^n\cdot mX^{m-1}.$$

On a aussi la formule pour la composition :

$$P(Q(X))'=P'(Q(X))Q'(X).$$

Par linéarité en $P(X)$ il suffit de montrer le cas $P(X)=X^n$, qui est $\big(Q(X)^n\big)'=nQ(X)^{n-1}Q'(X)$, qui se démontre par récurrence sur n via la formule de Leibniz.

Si le corps K est de caractéristique 0, les dérivées formelles se comportent comme on s'attend d'eux : par exemple : $\deg P'(X)=\deg P(X)-1$; $P'(X)=0$ ssi X constante, etc. Mais en caractéristique $p>0$ c'est différent. Par exemple, si K est de caractéristique 2, alors pour tout tout polynôme de la forme

$$P(X)=a_0+a_2X^2+a_4X^4+\cdots+a_{2m}X^{2m}=Q(X^2)$$

satisfait à $P'(X)=0$.

Définition. Un polynôme non nul $P(X)$ avec coefficients dans un corps L est *scindé sur L* s'il se factorise en $L[X]$ en facteurs de degré 1

$$P(X)=a(X-r_1)(X-r_2)\cdots(X-r_d).$$

avec $a,r_1,\ldots,r_d\in L$. Alors en groupant ensemble les r_i répétés on peut l'écrire aussi

$$P(X)=a(X-r_1)^{m_1}(X-r_2)^{m_2}\cdots(X-r_k)^{m_k}$$

avec $a \in K^\times$, les $r_1, \dots, r_k \in K$ distincts, et les $m_i \geq 1$. Alors m_i est la *multiplicité* de la racine r_i de $P(X)$. Une racine *simple* est une racine de multiplicité 1. Une racine *multiple* est de multiplicité ≥ 2.

La multiplicité m d'une racine $r \in K$ de $P(X) \in K[X]$ se caractérise aussi par une factorisation $P(X) = (X-r)^m Q(X)$ avec $Q(r) \neq 0$.

On a vu que pour tout polynôme non nul $P(X) \in K[X]$ il y a une extension L de K tels que $P(X)$ soit scindé sur L. La sous-extension $K(r_1, r_2, \dots, r_d)$ engendré par les racines est le *corps de décomposition* de $P(X)$ sur K.

Théorème 2.1. *Soit $P(X) \in K[X]$ non nul, et soit L une extension de K sur lequel $P(X)$ est scindé. Alors toutes les racines de $P(X)$ dans L sont simples si et seulement si* $\operatorname{pgcd}(P(X), P'(X)) = 1$ *dans $K[X]$.*

Le calcul essentiel est le lemme suivant.

Lemme 2.2. *Soit $P(X) \in L[X]$ non nul avec une racine $r \in L$. Alors r est une racine simple de $P(X)$ ssi $P'(r) \neq 0$.*

Preuve. On peut écrire $P(X) = (X-r)^m Q(X)$ avec $m \geq 1$ et $Q(r) \neq 0$. Quand $m = 1$ la dérivée est $P'(X) = Q(X) + (X-r)Q'(X)$, et $P'(r) = Q(r) \neq 0$. Quand $m \geq 2$ on a

$$P'(X) = m(X-r)^{m-1}Q(X) + (X-r)^m Q'(X).$$

avec $m - 1 \geq 1$, et $P'(r) = 0$. (La multiplicité de r comme racine de $P'(X)$ est $\geq m-1$ avec égalité ssi la caractéristique de L ne divise pas m.) □

Preuve. Les polynômes $P(X)$ et $P'(X)$ sont premiers entre eux dans $K[X]$ ssi ils sont premiers entre eux dans $L[X]$ parce que (par exemple) dans les calculs de l'algorithme d'Euclide les quotients et restes dans $K[X]$ sont des quotients et restes dans $L[X]$. Donc on peut supposer $K = L$ sans perte de généralité.

($\Rightarrow$) Si toutes les racines r_i de $P(X)$ sont simples, alors par le lemme $P'(r_i) \neq 0$ pour tout r_i, et aucun $X - r_i$ ne divise $P'(X)$. Comme les $X - r_i$ sont les seuls facteurs irréductibles de $P'(X)$, on voit que $P(X)$ et $P'(X)$ sont premiers entre eux.

($\Leftarrow$) Si $P(X)$ a une racine multiple r, alors on a $P'(r) = 0$ par le lemme. Donc $X - r$ est un diviseur commun de $P(X)$ et $P'(X)$, et ils ne sont pas premiers entre eux. □

Théorème 2.3. *Soit K un corps de caractéristique **nulle**, $P(X) \in K[X]$ un polynôme irréductible, et L une extension de K où $P(X)$ est scindé. Alors les racines de $P(X)$ dans L sont simples.*

Preuve. Le degré du polynôme irréductible est $d \geq 1$. En caractéristique 0 la dérivée $P'(X)$ est non nul de degré $d-1$. Donc il n'est pas divisible par $P(X)$. Mais quand un irréductible ne divise pas un autre polynôme, il est premier avec lui. □

Lemme 2.4. *Soit $P(X) = a(X-r_1)(X-r_2)\cdots(X-r_n)$ un polynôme scindé. Alors les valeurs de $P'(X)$ aux racines de $P(X)$ sont*

$$P'(r_i) = a \prod_{\substack{1 \leq j \leq n \\ j \neq i}} (r_i - r_j).$$