UFR Sciences
M2 Calcul Formel M. WEIMANN

Polynômes à plusieurs indéterminées

Références :

- Bostan, Chyzak, Giusti, Lebreton, Lecerf, Salvy, Schost, *Algorithmes efficaces en Calcul Formel* (2017).
- Cours de Calcul Formel de Michel Cretin *http ://ufr-mathematiques.univ-lyon1.fr/Agregation/*.
- Calcul mathématiques avec Sage (chapitre 9).
- Cox, Little, O'Shea, *Ideals, Varieties, and Algorithms* (en anglais, disponible en pdf sur le web). Dépasse le programme de l'agreg, mais accessible, beaucoup d'exos.

1 Systèmes polynomiaux à deux variables. Résultant.

On cherche à résoudre un système d'équations polynomiales $A(X,Y) = B(X,Y) = 0$ à deux variables. Le résultant va nous permettre d'éliminer la variable Y afin de déterminer un polynôme dont les racines donnent les abscisses des solutions du système.

1.1 Résultant

Inspiré de [BCGLSS]. Soient $A = a_m Y^m + \cdots + a_0$ et $B = b_n Y^n + \cdots + b_0$ dans $\mathbb{A}[Y]$ ($\mathbb{A}$ un anneau). On appelle matrice de Sylvester de A et B la matrice carrée de taille $m+n$ définie par

$$\mathrm{Syl}(A,B) = \begin{pmatrix} a_m & \cdots & a_1 & a_0 & & \\ & \ddots & & & \ddots & \\ & & a_m & a_{m-1} & \cdots & a_0 \\ b_n & \cdots & b_1 & b_0 & & \\ & \ddots & & & \ddots & \\ & & b_n & b_{n-1} & \cdots & b_0 \end{pmatrix}$$

avec les $n = \deg(B)$ premières lignes contenant les coefficients de A et les $m = \deg(A)$ lignes suivantes contenant les coefficients de B (ci-dessus représenté pour $n = m$).

Exercice 1 Dessiner la matrice de Sylvester pour $m = 4$ et $n = 2$.

La transposée de la matrice de Sylvester représente l'application linéaire

$$\begin{array}{rcl} \Phi : \mathbb{A}[Y]_{<n} \times \mathbb{A}[Y]_{<m} & \longrightarrow & \mathbb{A}[Y]_{<m+n} \\ (U\,,\,V) & \longmapsto & AU + BV \end{array}$$

dans les bases $(Y^{n-1},0),\ldots,(Y,0),(1,0),(0,Y^{m-1}),\cdots,(0,Y),(0,1)$ et $(Y^{m+n-1},\ldots,Y,1)$.

Définition 1 Le résultant de A et B est le déterminant de la matrice de Sylvester $\mathrm{Syl}(A,B)$. Il est noté $\mathrm{Res}(A,B)$, ou $\mathrm{Res}_Y(A,B)$ si l'on veut insister sur l'élimination de la variable Y.

Une propriété fondamentale du résultant est donnée par la proposition suivante, que l'on utilisera ensuite dans le cadre de l'élimination.

Proposition 1 Soient $A, B \in \mathbb{A}[Y]$. Si $\mathbb{A}$ est un corps, alors A et B sont premiers entre eux si et seulement si $\mathrm{Res}(A, B) \neq 0$.

Preuve. Les combinaisons linéaires des lignes de la matrice de Sylvester donnent les coefficients des polynômes qui sont dans l'image de Φ. Lorsque $\mathbb{A}$ est un corps, au vu de la relation de Bézout, le pgcd $D = \mathrm{pgcd}(A, B)$ est dans l'image, et il est caractérisé comme l'élément unitaire de plus petit degré qui peut l'être. En particulier, une base de l'image est donnée par $D, XD, \ldots, X^{m+n-\deg D}$, de dimension $m + n - \deg D$. Par le théorème du rang, il s'ensuit que le degré de D est égal à la dimension du noyau de $\mathrm{Syl}(A, B)$. D'où le résultat. □

Remarque 1 Cette preuve montre de plus que le degré de $D = \mathrm{pgcd}(A, B)$ est le nombre de lignes nulles de la forme échelonnée en ligne de $\mathrm{Syl}(A, B)$ et que les coefficients de D se lisent sur la dernière ligne non nulle de la forme échelonnée en ligne. La matrice de Sylvester est ainsi intimement liée à l'algorithme d'Euclide étendu.

La proposition suivante montre que le résultant obéit à une relation de type Bezout *à coefficients dans l'anneau* $\mathbb{A}$. C'est un point clé pour la résolution de systèmes polynomiaux.

Proposition 2 Soient $A, B \in \mathbb{A}[Y]$. Il existe $U, V \in \mathbb{A}[Y]$ avec $\deg U < \deg B$ et $\deg V < \deg A$, tels que $\mathrm{Res}(A, B) = UA + VB$.

Preuve. On pense la matrice à coefficients dans l'anneau $\mathbb{A}[Y]$. En ajoutant à la dernière colonne de la matrice de Sylvester la colonne i multipliée par Y^{n+m-i} pour $i = 1, \ldots, n+m$, on fait apparaître dans la dernière colonne les $n+m$ polynômes $Y^{n-1}A, \ldots, YA, A$, et $Y^{m-1}B, \ldots, YB, B$, sans avoir changé le déterminant. Le développement du déterminant par rapport à la dernière colonne permet de conclure. □

D'autres propriétés élémentaires du résultant (déjà vues en M1) sont rappelées en fin de section (formule de Poisson, spécialisation). Pour l'instant, revenons à nos moutons : résoudre un système de deux équations polynomiales à deux inconnues.

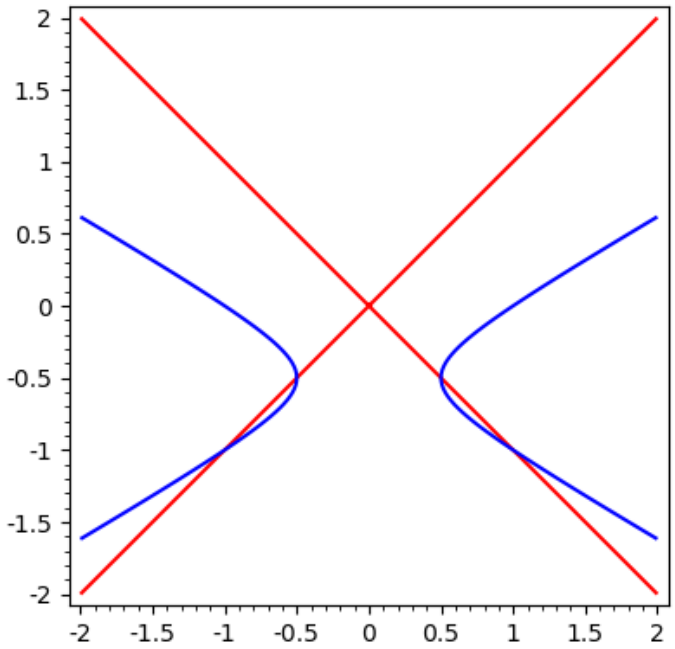
1.2 Elimination via le résultant

Considérons le système polynomial

$$\begin{cases} A = -3Y^2 - 3Y + X^2 - 1 = 0 \\ B = -Y^2 + X^2 = 0 \end{cases}$$

Les polynômes A et B définissent deux courbes dans $\mathbb{R}^2$ (ou $\mathbb{C}^2$). L'intersection de ces courbes définit un ensemble fini de points, dont les coordonnées se calculent ici aisément :

$$(-1, -1,), (1, -1), (-1/2, -1/2), (1/2, -1/2)$$



Le résultant va nous permettre de déterminer les abscisses α de ces points d'intersections dans le cas général en *éliminant* la variable Y. Déterminer les ordonnées de ces points reviendra alors à calculer les racines communes des polynômes univariés $A(\alpha, Y)$ et $B(\alpha, Y)$ (i.e. racine du pgcd de ces deux polynômes).

Soit $\mathbb{K}$ un corps. Un polynôme à deux variables X, Y à coefficients dans le corps $\mathbb{K}$ peut être vu comme un polynôme à une variable Y à coefficients dans l'anneau $\mathbb{K}[X]$. En particulier, cela fait sens de considérer $\mathrm{Res}_Y(A, B) \in \mathbb{K}[X]$ pour $A, B \in \mathbb{K}[X, Y]$. Le résultat fondamental pour notre propos est le suivant :

Proposition 3 Soient $A, B \in \mathbb{K}[X, Y]$. Notons a_m et b_n les coefficients de tête de A et B. Alors les racines de $\mathrm{Res}_Y(A, B) \in \mathbb{K}[X]$ sont d'une part les abscisses des solutions du système $A = B = 0$ et d'autre part les racines communes des coefficients dominants $a_m, b_n \in \mathbb{K}[X]$. Les racines sont considérées ici dans une clôture algébrique de $\mathbb{K}$.

Preuve. Les racines de a_m et b_n sont des zéros du résultant d'après la formule de Poisson (Théorème 1). La Proposition 2 fournit une identité type Bezout

$$\mathrm{Res}_Y(A, B) = AU + BV, \qquad U, V \in \mathbb{K}[X, Y].$$

Si (α, β) est solution du système $A = B = 0$, alors en évaluant la relation de Bezout en (α, β), on voit que $\mathrm{Res}_Y(A, B)(\alpha) = 0$. Réciproquement, si α est racine de $\mathrm{Res}_Y(A, B)$, la Proposition 4 appliquée au morphisme unitaire d'évaluation $X \mapsto \alpha$ assure que $\mathrm{Res}_Y(A(\alpha, Y), B(\alpha, Y)) = 0$. Si α n'est pas racine des coefficients de tête a_m et b_n, la Proposition 1 (dans le corps $\mathbb{A} = \mathbb{K}(\alpha)$) assure que $A(\alpha, Y)$ et $B(\alpha, Y)$ ont un pgcd non trivial, i.e. il existe β tel que $A(\alpha, \beta) = B(\alpha, \beta) = 0$. $\square$.

Moralement, les racines communes de a_m et b_n correspondent aux abscisses des "points d'intersections à l'infini" des courbes $A = 0$ et $B = 0$ (correspondant aux branches infinies ayant même asymptotes). Une fois calculé $R(X) = \mathrm{Res}_Y(A, B)$, il ne reste plus alors qu'à calculer les solutions de $\mathrm{pgcd}(A(\alpha, Y), B(\alpha, Y))$ où α parcourt les racines de R qui ne sont pas racines communes de a_m et b_n.

Exemple 1. Reprenant l'exemple précédent, $A = -3Y^2 - 3Y + X^2 - 1 = 0$ et $B = -Y^2 + X^2 = 0$, on trouve

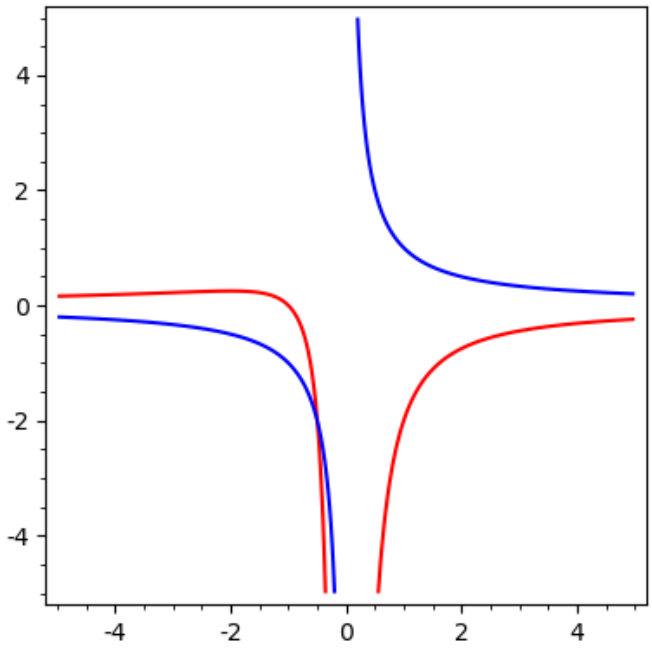
$$\mathrm{Res}_Y(A,B) = \det \begin{pmatrix} -3 & -3 & X^2-1 & 0 \\ 0 & -3 & -3 & X^2-1 \\ -1 & 0 & X^2 & 0 \\ 0 & -1 & 0 & X^2 \end{pmatrix} = 4X^4 - 5X^2 + 1 = 4(X^2-1)(X^2 - \frac{1}{4})$$

On retrouve bien le fait que les abscisses des solutions du système sont ± 1 et $\pm -1/2$. Les solutions du système s'en déduisent aisément.

Exemple 2. Considérons cette fois $A = X^2Y + X + 1$ (courbe rouge ci-dessous) et $B = XY - 1$ (courbe bleue). On trouve

$$\mathrm{Res}_Y(A,B) = -X(2X+1)$$

La racine $\alpha = 0$ du résultant correspond à l'asymptote verticale $X = 0$ (zéro commun des coefficients dominants de A et B), et la racine $\alpha = -\frac{1}{2}$ correspond à la vraie solution $(-\frac{1}{2}, -2)$.



Remarque 2 Calculer les abscisses des points d'intersection revient à calculer la projection de ces points de $\mathbb{R}^2$ sur la droite $\mathbb{R}$ d'équation $Y = 0$. Autrement dit : "élimination (algèbre) signifie projection (géométrie)". Ceci se traduit algébriquement par le calcul de l'idéal principal

$$I = (A,B) \cap \mathbb{K}[X].$$

Les résultats précédents assurent que $\mathrm{Res}_Y(A,B) \in I$ et, encore mieux, le radical de I est engendré par le radical du résultant (factorisation sans carrés). Attention, le résultant n'engendre pas nécessairement I.

1.3 Implicitation

Soit $\mathbb{K}$ un corps algébriquement clos. Il est difficile de déterminer si un point donné vit sur une courbe paramétrée

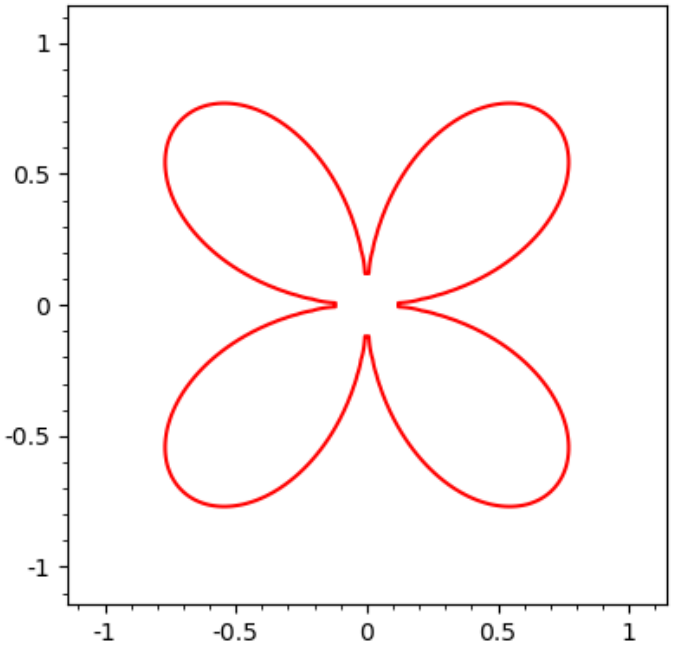
$$X = A(T), \quad Y = B(T), \qquad A, B \in \mathbb{K}(T).$$

Pour ce genre de problème, il est souhaitable de calculer une équation implicite de cette courbe, i.e. un polynôme $R \in \mathbb{K}[X,Y]$ qui s'annule sur la courbe.

Un analogue à 3 variables de la Proposition 3 assure qu'il suffit de calculer le résultant en T des numérateurs $P, Q \in \mathbb{K}[X,Y,T]$ de $X - A(T)$ et $Y - B(T)$. En effet $\mathrm{Res}_T(P,Q)$ est alors un polynôme en (X,Y) qui s'annule en (α,β) si et seulement si $P(\alpha,\beta,T)$ et $Q(\alpha,\beta,T)$ ont un zéro commun, i.e. ssi il existe $t \in \mathbb{K}$ tel que $\alpha = A(t)$ et $\beta = B(t)$, i.e. ssi (α,β) vit sur la courbe paramétrée (certaines valeurs de α,β correspondent éventuellement à des points "à l'infini", i.e. à des valeurs de t en lesquelles les dénominateurs de A ou B s'annulent.

Exemple (quadrifolium). Considérons la courbe paramétrée par

$$A(T) = \frac{4\left(T^2-1\right)^2 T}{\left(T^2+1\right)^3}, \qquad B(T) = \frac{8\left(1-T^2\right)T^2}{\left(T^2+1\right)^3}$$



(les feuilles sont censées s'intersecter à l'origine, ce que le dessin représente mal). Une équation implicite est donc donnée par $R(X,Y) = 0$ où

$$R(X,Y) := \mathrm{Res}_T\left(X\left(T^2+1\right)^3 - 4\left(T^2-1\right)^2, Y\left(T^2+1\right)^3 - 8\left(T^2-1\right)T^2\right).$$

On trouve ici $R = X^6 + 3X^4Y^2 + 3X^2Y^4 + Y^6 - 4X^2Y^2$ à une constante multiplicative près.

1.4 Propriétés du résultant

Théorème 1 (Formule de Poisson) Soient $A = a(Y-\alpha_1)\cdots(Y-\alpha_m)$ et $B = b(Y-\beta_1)\cdots(Y-\beta_n)$. On a

$$\mathrm{Res}_Y(A,B) = a^n b^m \prod_{i,j}(\alpha_i - \beta_j) = a^n \prod_i B(\alpha_i) = (-1)^{mn} b^m \prod_j A(\beta_j) = (-1)^{mn}\,\mathrm{Res}_Y(B,A).$$

Preuve (esquisse). Le facteur ab^n vient de la multiplicativité du déterminant. On peut donc supposer $a = b = 1$. Si $\alpha_i = \beta_j$, alors A et B ne sont pas copremiers et le résultant est nul. Ainsi, le produit des $\alpha_i - \beta_j$ divise $\mathrm{Res}_Y(A,B)$ vu comme élément de $\mathbb{Z}[\alpha_1,\ldots,\beta_m]$. Par ailleurs, le degré en α_i des

coefficients des , n premières lignes de la matrice de Sylvester est au plus 1, et ce degéré est nul pour les autres lignes. Donc le résultant est de degré au plus n en chaque α_i. De même il est de degré au plus n en chaque β_j. Donc $\mathrm{Res}_Y(A,B)$ est une constante fois le produit des $\alpha_i - \beta_j$. En spécialisant $A = Y^n$, la matrice de Sylvester est triangulaire de déterminant $B(0)^m$, ce qui donne le facteur 1 et conclut la preuve. □.

Exercice 2 Retrouver les résultants des exemples 1 et 2 à l'aide de la formule de Poisson.

Corollaire 1 (Multiplicativité) Pour tout $A, B, C \in \mathbb{A}[Y]$, on a

$$\mathrm{Res}(AB, C) = \mathrm{Res}(A, C)\,\mathrm{Res}(B, C)$$

Corollaire 2 (Résultant et algorithme d'Euclide) . Soit $\mathbb{K}$ un corps et soit $A = QB + R$ la division euclidienne de A par B dans $\mathbb{K}[Y]$. Soit $r = \deg(R)$. On a :

$$\mathrm{Res}(A, B) = (-1)^{mn} b^{m-r}\,\mathrm{Res}(B, R)$$

Corollaire 3 (Calcul rapide du résultant) Le résultant peut se calculer avec un algorithme de type Euclide étendu avec $O(mn)$ opérations dans $\mathbb{K}$, ou $O(M(n)\log(n))$ opérations en adaptant Euclide étendu rapide (en supposant ici $n \geq m$).

Exercice 3 Ecrire un tel algorithme et retrouver les résultants des exemples 1 et 2.

Remarque 3 Ce dernier corollaire assure que si $A, B \in \mathbb{K}[X, Y]$, alors $\mathrm{Res}_Y(A, B) = UA + VB$ avec $U, V \in \mathbb{K}(X)$. La proposition 2 assure que l'on peut en fait choisir U, V dans l'anneau $\mathbb{K}[X]$.

On renvoie à [BCGLLSS] pour les deux résultats suivants.

Proposition 4 (Spécialisation du résultant) Soit $\Phi : \mathbb{A} \to \mathbb{B}$ un morphisme d'anneau unitaires que l'on étend en $\Phi : \mathbb{A}[Y] \to \mathbb{B}[Y]$ coefficients par coefficients. Si $A \in \mathbb{A}[Y]$ est *unitaire*, alors

$$\phi(\mathrm{Res}(A, B)) = \mathrm{Res}(\phi(A), \phi(B)).$$

Exemple. Suppose $\mathbb{A} = \mathbb{K}[X]$ et $\mathbb{B} = \mathbb{K}$ et soit $\Phi : \mathbb{K}[X] \to \mathbb{K}$ le morphisme d'évaluation en $X = 0$. Soit $A = XY^2 + X + Y$ et $B = 2Y - 1$. On a

$$\mathrm{Res}_Y(A, B) = 5X - 2, \qquad \mathrm{Res}_Y(A, B)(0) = -2 \neq \mathrm{Res}_Y(A(0, Y), B(0, Y)) = 1.$$

Résultant et spécialisation ne commutent pas ici du fait que A n'est pas unitaire (en Y).

Proposition 5 Soient $A, B \in \mathbb{A}[Y]$ avec A unitaire. Alors $\mathrm{Res}(A, B)$ est le déterminant de l'endomorphisme de multiplication par B dans la $\mathbb{A}$-algèbre $\mathbb{A}[Y]/(A)$.

2 Polynômes symétriques à n indéterminées

On démontre le théorème fondamental des fonctions symétriques : tout polynôme symétrique est un polynôme en les fonctions symétriques élémentaires. La preuve est constructive. Elle permet d'introduire l'ordre monomial lexicographique et d'avoir une première idée de la division à plusieurs variables. On finit avec les identités de Newton (facultatif).

2.1 Polynômes à n indéterminées. Terminologie.

Soit $\mathbb{K}$ un corps. On note $\mathbb{K}[X_1, \ldots, X_n]$ l'ensemble des polynômes en les indéterminées $X_1, \ldots, X_n$. C'est un anneau commutatif, dont les éléments sont les sommes finies de la forme

$$P(X_1, \ldots, X_n) = \sum_{\alpha_1, \ldots, \alpha_n \in \mathbb{N}} c_{\alpha_1, \ldots, \alpha_n} X_1^{\alpha_1} \cdots X_n^{\alpha_n}.$$

On notera parfois plus brièvement $X = (X_1, \ldots, X_n)$ et $\alpha = (\alpha_1, \ldots, \alpha_n)$ de sorte que

$$P(X) = \sum_{\alpha \subset \mathbb{N}^n} c_\alpha X^\alpha,$$

où il est entendu que $X^\alpha = X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. Les c_α sont les *coefficients* de P. Les X^α sont les *monômes* de P. Les $c_\alpha X^\alpha$ sont les *termes* de P.

Le *degré* d'un monôme X^α est $|\alpha| := \alpha_1 + \cdots + \alpha_n$. Le degré d'un polynôme est le max des degrés de ses monômes. On utilise en général la convention $\deg(0) = -\infty$.

On dit qu'un polynôme P est *homogène* si tous ses monômes ont même degrés. Les polynômes homogènes de degré 1 sont appelés des formes linéaires, de la forme

$$P(X) = c_1 X_1 + \cdots + c_n X_n$$

Les polynômes homogènes de degré 2 sont appelés des formes quadratiques, de la forme

$$P(X) = c_{11} X_1^2 + c_{12} X_1 X_2 + \cdots + c_{ij} X_i X_j + \cdots + c_{nn} X_n^2.$$

Exercice 4 L'ensemble R_d des polynômes de degré au plus d forme un sous-espace vectoriel de $\mathbb{K}[X]$ de dimension $\binom{n+d}{d}$. L'ensemble H_d des polynômes *homogènes* de degré d forme un sous-espace vectoriel de $\mathbb{K}[X]$ de dimension $\binom{n+d-1}{d}$. On a l'égalité $R_d = H_0 \oplus H_1 \oplus \cdots \oplus H_d$.

Exercice 5

1. Montrer que P est homogène de degré d si et seulement si

$$P(TX_1, \ldots, TX_n) = T^d P(X_1, \ldots, \ldots, X_n).$$

2. (*Identité d'Euler*) Montrer que si P est homogène de degré d, alors

$$dP = \sum_{i=1}^{n} X_i \frac{\partial P}{\partial X_i}.$$

2.2 Théorème fondamental des polynômes symétriques

Définition 2 On dit que $P \in \mathbb{K}[X_1, \ldots, X_n]$ est un polynôme symétrique s'il est invariant par permutation des variables, c'est à dire si

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

pour toute permutation $\rho \in S_n$.

Avec les notations de l'introduction, cela est équivalent au fait que

$$c_{\alpha_1, \ldots, \alpha_n} = c_{\rho(\alpha_1), \ldots, \rho(\alpha_n)} \quad \forall\, \rho \in S_n.$$

Par exemple, le polynôme $X_1^2 + 4X_1X_2 + X_2^2$ est un polynôme symétrique de $\mathbb{K}[X_1, X_2]$.

Définition 3 Pour $k = 1, \ldots, n$, le k-ème polynôme symétrique élémentaire en n variables est

$$\sigma_k(X_1, \ldots, X_n) = \sum_{1 \leq \alpha_1 < \cdots < \alpha_k \leq n} X_{\alpha_1} \cdots X_{\alpha_k}.$$

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

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

Théorème 2 (Relations racines et polynômes symétriques élémentaires) Soient $r_1, \ldots, r_n \in \mathbb{K}$. Notons $\sigma_i := \sigma_i(r_1, \ldots, r_n)$. On a la relation :

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

Exercice 6 Prouver ce théorème en utilisant la formule (que l'on démontre par récurrence)

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

Il est clair que les σ_i sont des polynômes symétriques, et, plus généralement tout polynôme en les σ_i est un polynôme symétrique. Il se trouve que la réciproque est vraie :

Théorème 3 (Théorème fondamental des fonctions symétriques) Tout polynôme symétrique $P \in \mathbb{K}[X_1, \ldots, X_n]$ s'écrit de manière unique comme polynôme en $\sigma_1, \ldots, \sigma_n$.

Par exemple, dans $\mathbb{K}[X_1, X_2, X_3]$, on a la relation

$$X_1^2 + X_2^2 + X_3^2 = \sigma_1^2 - 2\sigma_2$$

2.3 Preuve du théorème fondamental

L'ordre lexicographique. On ordonne les monômes en munissant l'ensemble $\mathbb{N}^n$ des multi-degrés $\alpha = (\alpha_1, \ldots, \alpha_n)$ de l'ordre total lexicographique. Autrement dit :

$$X_1^{\alpha_1} \cdots X_n^{\alpha_n} \succ X_1^{\beta_1} \cdots X_n^{\beta_n}$$

si et seulement si, arrivé au premier k tel que $\alpha_k \neq \beta_k$, on a $\alpha_k > \beta_k$. Par exemple $X_1X_3^4 \prec X_1X_2$. Noter aussi que $X_1 \succ X_2 \succ \cdots \succ X_n \succ 1$. Les monômes forment ainsi un *ensemble bien ordonné* (ordre *total* pour lequel toute partie non vide a un plus petit élément). En particulier, le principe de récurrence est valable (récurrences transfinies). Cet ordre s'étend en un ordre (partiel cette fois) sur les termes $c_\alpha X^\alpha$.

Tout polynôme P admet un unique monôme maximal X^α pour cet ordre. On l'appelle le monôme dominant de P, noté en général $\mathrm{lm}(P)$ (leading monomial). Le terme $c_\alpha X^\alpha$ de P correspondant s'appelle le terme dominant de P, noté $\mathrm{lt}(P)$ (leading term).

Par exemple, $\mathrm{lm}(2X_1X_3 - 5X_2^3X_3^4) = X_1X_3$ et $\mathrm{lt}(2X_1X_3 - 5X_2^3X_3^4) = 2X_1X_3$.

Lemme 1 Soient $P, Q \in \mathbb{K}[X_1, \ldots, X_n]$. On a

$$\mathrm{lt}(PQ) = \mathrm{lt}(P)\,\mathrm{lt}(Q) \qquad \text{et} \qquad \mathrm{lt}(P - Q) \preceq \max(\mathrm{lt}(P), \mathrm{lt}(Q))$$

avec inégalité stricte si et seulement si $\mathrm{lt}(P) = \mathrm{lt}(Q)$.

Exercice 7 Prouver le lemme.

Existence. On montre la partie existence du théorème par induction . Soit P un polynôme symétrique. Supposons inductivement que tout polynôme symétrique non nul de terme dominant strictement inférieur à celui de P soit exprimable en fonction polynomiale des polynômes symétriques élémentaires, et montrons qu'il en est de même pour P.

Soit $cX_1^{\alpha_1}\cdots X_n^{\alpha_n} = \mathrm{lt}(P)$. Par hypothèse, les monômes $cX_1^{\rho(\alpha_1)}\cdots X_n^{\rho(\alpha_n)}$ sont aussi des monômes de P et on déduit que $\alpha_1 \geq \cdots \geq \alpha_n$ par maximalité de $(\alpha_1,\ldots,\alpha_n)$. Les entiers définis par $t_k := \alpha_k - \alpha_{k+1}$ si $k < n$ et $t_n = \alpha_n$ sont donc positifs ou nuls. Le polynôme symétrique

$$Q = c\sigma_1^{t_1}\cdots\sigma_n^{t_n}$$

a même terme dominant que P :

$$\mathrm{lt}(Q) = cX_1^{t_1}(X_1X_2)^{t_2}\cdots(X_1\cdots X_n)^{t_n} = cX_1^{\alpha_1}\cdots X_n^{\alpha_n} = \mathrm{lt}(P)$$

(la première égalité se déduisant du Lemme 1). Le polynôme symétrique $P - Q$ a donc un terme dominant strictement plus petit que celui de P et on conclut par induction que $P - Q$ est un polynôme en les σ_i. Donc P l'est aussi.

Unicité. Pour l'unicité, il suffit de montrer que si $T \in \mathbb{K}[S_1,\ldots,S_n]$ en non nul, alors $P(X) = T(\sigma_1,\ldots,\sigma_n) \in \mathbb{K}[X_1,\ldots,X_n]$ est non nul. Soit donc T un tel polynôme non nul. Soit $cS_1^{t_1}\cdots S_n^{t_n}$ le monôme de T pour lequel le monôme $cX_1^{t_1}(X_1X_2)^{t_2}\cdots(X_1\cdots X_n)^{t_n}$ est maximal. D'après ce qui précède, ce monôme apparaît dans P avec un coefficient non nul. Donc P est non nul. □.

Algorithme. On déduit de cette preuve constructive l'algorithme suivant permettant d'exprimer un polynôme symétrique P comme $P = Q(\sigma_1,\ldots,\sigma_n)$ pour un certain polynôme $Q \in \mathbb{K}[S_1,\ldots,S_n]$ que l'on cherche à calculer. On pose $P_0 = P$ et $Q_0 = 0$. Supposons que $P_i \in \mathbb{K}[X_1,\ldots,X_n]$ et $Q_i \in \mathbb{K}[S_1,\ldots,S_n]$ aient été construits, vérifiant

$$P_i = P - Q_i(\sigma_1,\ldots,\sigma_n) \in \mathbb{K}[X_1,\ldots,X_n]$$

Si $P_i = 0$, on retourne $Q = Q_i$. Sinon, on calcule $\mathrm{lt}(P_i) = cX_1^{\alpha_1}\cdots X_n^{\alpha_n}$ et on pose

$$Q_{i+1} = Q_i + cS_1^{t_1}\cdots S_n^{t_n} \quad et \quad P_{i+1} = P_i - c\sigma_1^{t_1}\cdots\sigma_n^{t_n}$$

où $t_k = \alpha_k - \alpha_{k+1} \geq 0$ pour $k = 1,\ldots,n-1$ et $t_n = \alpha_n$ (cf ci-dessus). D'après ce qui précède, le terme dominant de P_{i+1} est strictement inférieur à celui de P_i, donc il existe un plus petit indice m tel que $P_m = 0$. Le polynôme cherché est $Q = Q_m$.

Exercice 8 Dérouler l'algorithme pour exprimer le polynôme

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

comme un polynôme en les fonctions symétriques élémentaires $\sigma_1, \sigma_2, \sigma_3$.

2.4 Identités de Newton

Soit $k \in \mathbb{N}$. On appelle k-ème somme de Newton de $\mathbb{K}[X_1,\ldots,X_n]$ le polynôme

$$s_k = X_1^k + \cdots + X_n^k.$$

Les sommes de Newton sont des polynômes symétriques. Elles sont liées aux fonctions symétriques élémentaires par le résultat suivant :

Proposition 6 (Identités de Newton) On pose par convention $\sigma_0 = 1$ et $\sigma_i = 0$ si $i > n$ ou $i < 0$. Pour tout $k \geq 1$, on a l'égalité :

$$\sigma_0 s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k = 0.$$

Exercice 9 Notons $\sigma_k^{(n)}$ le k-ème polynôme symétrique élémentaire en n indéterminées. Montrer que

$$\sigma_k^{(n)} = \sigma_k^{(n-1)} + \sigma_{k-1}^{(n-1)} X_n.$$

En combinant avec l'égalité $s_k^{(n)} = s_k^{(n-1)} + X_n^k$, en déduire une preuve de la proposition par récurrence sur le nombre n de variables.

Théorème 4 Si $\mathbb{K}$ est de caractéristique 0, alors σ_k est un polynôme en $s_1, \ldots, s_k$ pour tout $k = 1, \ldots, n$. Il s'ensuit que tout polynôme symétrique est un polynôme en $s_1, \ldots, s_n$.

Preuve. Si $\mathbb{K}$ est de caractéristique zéro, alors k est inversible dans $\mathbb{K}$. Il suit alors immédiatement (par induction sur k) de la proposition précédente que σ_k est un polynôme en $s_1, \ldots, s_k$. La dernière assertion découle du théorème fondamental des fonctions symétriques. □

Exercice 10 Montrer que si $\mathbb{K}$ est de caractéristique 2, il est impossible d'écrire $X_1X_2 \in \mathbb{K}[X_1, X_2]$ comme polynôme en s_1 et s_2.

3 Division des polynômes à plusieurs variables

3.1 Bon ordre monomial

Notons $\mathcal{M}_n$ l'ensemble des monômes de $\mathbb{K}[X_1, \ldots, X_n]$. Un *bon ordre* sur $\mathcal{M}_n$ est une relation d'ordre sur les monômes qui est totale (on a soit $X^\alpha \prec X^\beta$, soit $X^\alpha = X^\beta$, soit $X^\alpha \succ X^\beta$) qui vérifie de plus

$$X^\alpha \prec X^\beta \Longrightarrow X^{\alpha+\gamma} \prec X^{\beta+\gamma} \quad \text{(ordre monomial)} \qquad \text{et} \qquad 1 \preceq X^\alpha \quad \text{(ordre admissible)}$$

La seconde condition assure que toute partie non vide de $\mathcal{M}_n$ a un plus petit élément pour un bon ordre. Notons aussi que l'on a

$$X^\alpha | X^\beta \Longrightarrow X^\alpha \preceq X^\beta$$

ce qui fait d'un ordre monomial une relation d'ordre plus fine que l'ordre partiel donné par la divisibilité (ces deux notions coïncidant dans le cas d'une seule variable).

Principaux ordres monomiaux, avec l'exemple de $\mathbb{Q}[x,y,z]$	
`lex`	$\boldsymbol{x^\alpha} < \boldsymbol{x^\beta} \iff \alpha_1 < \beta_1$ ou $(\alpha_1 = \beta_1$ et $\alpha_2 < \beta_2)$ ou $\ldots$ ou $(\alpha_1 = \beta_1, \ldots, \alpha_{n-1} = \beta_{n-1}$ et $\alpha_n < \beta_n)$ $x^3 > x^2y > x^2z > x^2 > xy^2 > xyz > xy > xz^2 > xz > x > y^3 > y^2z > y^2 > yz^2 > yz > y > z^3 > z^2 > z > 1$
`invlex`	$\boldsymbol{x^\alpha} < \boldsymbol{x^\beta} \iff \alpha_n < \beta_n$ ou $(\alpha_n = \beta_n$ et $\alpha_{n-1} < \beta_{n-1})$ ou $\ldots$ ou $(\alpha_n = \beta_n, \ldots, \alpha_2 = \beta_2$ et $\alpha_1 < \beta_1)$ $z^3 > yz^2 > xz^2 > z^2 > y^2z > xyz > yz > x^2z > xz > z > y^3 > xy^2 > y^2 > x^2y > xy > y > x^3 > x^2 > x > 1$
`deglex`	$\boldsymbol{x^\alpha} < \boldsymbol{x^\beta} \iff \lvert\boldsymbol{\alpha}\rvert < \lvert\boldsymbol{\beta}\rvert$ ou $(\lvert\boldsymbol{\alpha}\rvert = \lvert\boldsymbol{\beta}\rvert$ et $\boldsymbol{x^\alpha} <_{\texttt{lex}} \boldsymbol{x^\beta})$ $x^3 > x^2y > x^2z > xy^2 > xyz > xz^2 > y^3 > y^2z > yz^2 > z^3 > x^2 > xy > xz > y^2 > yz > z^2 > x > y > z > 1$
`degrevlex`	$\boldsymbol{x^\alpha} < \boldsymbol{x^\beta} \iff \lvert\boldsymbol{\alpha}\rvert < \lvert\boldsymbol{\beta}\rvert$ ou $(\lvert\boldsymbol{\alpha}\rvert = \lvert\boldsymbol{\beta}\rvert$ et $\boldsymbol{x^\alpha} >_{\texttt{invlex}} \boldsymbol{x^\beta})$ $x^3 > x^2y > xy^2 > y^3 > x^2z > xyz > y^2z > xz^2 > yz^2 > z^3 > x^2 > xy > y^2 > xz > yz > z^2 > x > y > z > 1$

Contrairement à l'ordre partiel de divisibilité, les bons ordres monomiaux ont l'avantage de permettre de généraliser la division euclidienne au cas multivarié. Tout ordre monomial s'étend naturellement en un ordre (partiel) sur les termes $c_\alpha X^\alpha$.

On se fixe un bon ordre et on note :

— $\mathrm{lt}(P)$ le terme dominant (leading term) de P pour l'ordre considéré.

— $\mathrm{lm}(P)$ le monôme dominant (leading monomial) de P pour l'ordre considéré.

Le Lemme 1 reste valable pour tout bon ordre monomial.

3.2 Algorithme de division à plusieurs variables

Soient $G_1, \ldots, G_r \in \mathbb{K}[X_1, \ldots, X_n]$ une liste de "diviseurs" et soit $\prec$ un bon ordre monomial.

Définition 4 (Division multivariée) Une division multivariée de P par $G_1, \ldots, G_r$ (relativement à l'ordre $\prec$) consiste en une liste de quotients $Q_1, \ldots, Q_r$ et un reste R *en général non uniques* tels que

1. $P = Q_1G_1 + \cdots + Q_rG_r + R$
2. $\mathrm{lt}(G_i)$ ne divise aucun monôme de R pour tout i.
3. $\mathrm{lt}(Q_iG_i) \preceq \mathrm{lt}(P)$ pour tout i.

On dit que R est *une réduction* de P modulo l'idéal $\langle G_1, \ldots, G_r\rangle$.

L'algorithme est le suivant, très proche de la division euclidienne classique .

1. On initialise $Q_i = 0$ pour tout i et $R = 0$. On pose $F = P$.
2. Tant que $F \neq 0$:

— S'il existe i tel que $\mathrm{lt}(G_i)$ divise $\mathrm{lt}(F)$, alors pour le plus petit tel i,

$$Q_i \leftarrow Q_i + \frac{\mathrm{lt}(F)}{\mathrm{lt}(G_i)}, \qquad F \leftarrow F - \frac{\mathrm{lt}(F)}{\mathrm{lt}(G_i)}G_i$$

— Sinon :

$$R \leftarrow R + \mathrm{lt}(F), \qquad F \leftarrow F - \mathrm{lt}(F)$$

3. On retourne $Q_1, \ldots, Q_r, R$.

Théorème 5 L'algorithme termine est retourne une division multivariée de P par $Q_1, \ldots, Q_r$.

Preuve (esquisse). On vérifie à l'aide du lemme 1 que $\mathrm{lt}(F)$ est strictement décroissant (exo). Comme l'ordre lexicographique est un bon ordre (toute partie non vide admet un plus petit élément) on atteint $F = 0$ après un nombre fini d'étape et l'algorithme termine. Pour montrer que l'algorithme est correct, on vérifie d'abord (exo) qu'à toute étape de l'algorithme on a un "invariant de boucle"

$$F + \sum_{i=1}^{r} Q_i G_i + R = P.$$

Donc la condition 1 de la division multivariée est vérifiée. Puisque l'on ajoute à R des monômes non divisibles par $\mathrm{lt}(G_i)$, R est bien réduit modulo $(G_1, \ldots, G_r)$. Donc la condition 2 est vérifiée. Reste à montrer que $\mathrm{lt}(Q_i G_i) \preceq \mathrm{lt}(P)$. On le montre par induction sur le nombre d'étapes de l'algorithme. C'est vrai à l'initialisation. Si Q_i n'est pas modifié, aucun Q_j ne bouge, rien à montrer. Sinon, Q_j ne bouge pas pour $j \neq i$, mais $Q_i G_i$ est remplacé par $Q_i G_i + \frac{\mathrm{lt}(F)}{\mathrm{lt}(G_i)} G_i$ tandis que F devient $F - \frac{\mathrm{lt}(F)}{\mathrm{lt}(G_i)} G_i$. A l'aide du lemme 1, on déduit alors

$$\mathrm{lt}\left(Q_i G_i + \frac{\mathrm{lt}(F)}{\mathrm{lt}(G_i)} G_i\right) \preceq \max(\mathrm{lt}(Q_i G_i), \mathrm{lt}(F)) \preceq \max(\mathrm{lt}(P), \mathrm{lt}(F)) = \mathrm{lt}(P),$$

la deuxième inégalité par induction, et la première inégalité et l'égalité du fait de la stricte décroissance de $\mathrm{lt}(F)$. D'où le résultat. □.

La division multivariée par l'idéal $(G_1, \ldots, G_r)$ n'est pas bien définie en ce sens que le reste R va dépendre en général du choix des générateur ou de l'ordre des G_i. Afin de palier à cette difficulté, on introduit des systèmes de générateurs particuliers, les bases de Gröbner.

4 Bases de Gröbner

Les bases de Gröbner ont de nombreuses applications, au coeur de la résolution des systèmes polynomiaux. Moralement, elles permettent de ramener l'étude d'un idéal polynomial à l'étude d'un idéal monomial (i.e. engendré par des monômes), plus facile à appréhender.

On fixe un bon ordre monomial $\prec$ sur l'ensemble $\mathcal{M}_n$ des monômes $\mathbb{K}[X_1, \ldots, X_n]$.

4.1 Définition

Etant donné un idéal $I \subset \mathbb{K}[X_1, \ldots, X_n]$, on note $\mathrm{lt}(I)$ l'idéal engendré par les termes dominants des éléments de I. C'est donc un idéal monomial (engendré par des monômes). En particulier, $P \in \mathrm{lt}(I)$ si et seulement si chaque monôme de P est divisible par au moins l'un des monômes générateurs (exo).

Si $I = \langle G_1, \ldots, G_r \rangle$, on a l'inclusion évidente

$$\langle \mathrm{lt}(G_1), \ldots, \mathrm{lt}(G_r) \rangle \subset \mathrm{lt}(I),$$

mais ce n'est pas une égalité en général. Par exemple, si $G_1 = XY + 1$ et $G_2 = X^2 - 1$, on remarque que

$$YG_1 - XG_2 = X + Y \in I = \langle G_1, G_2 \rangle$$

de sorte que (avec l'ordre lexico) $X = \mathrm{lt}(X + Y) \in \mathrm{lt}(I)$. Pour autant, X n'appartient pas à l'idéal $\langle \mathrm{lt}(G_1), \mathrm{lt}(G_2) \rangle = \langle XY, Y^2 \rangle$.

Définition 5 Soit $I \subset \mathbb{K}[X_1, \ldots, X_n]$ un idéal. On dit qu'un système de générateur $(G_1, \ldots, G_r)$ de I est une base de Gröbner ou base standard (pour un bon ordre donné) s'il y a égalité

$$\langle \mathrm{lt}(G_1), \ldots, \mathrm{lt}(G_r) \rangle = \mathrm{lt}(I).$$

Autrement dit, le terme (ou monôme) dominant de tout $f \in I$ est divisible par le terme (ou monôme) dominant de l'un des G_i.

4.2 Division par une base de Gröbner

Proposition 7 Soit $I \subset \mathbb{K}[X_1, \ldots, X_n]$ un idéal. On a $P \in I$ si et seulement si son reste dans la division multivariée par une base de Gröbner $(G_1, \ldots, G_r)$ de I est nul.

Preuve. Si le reste est nul, on a $P \in I$ (Definition 4). Si $P \in I$, son reste est dans I. S'il est non nul, son monôme dominant est donc dans $\mathrm{lt}(I)$, donc multiple d'un des $\mathrm{lt}(G_i)$ par définition d'une base de Gröbner, ce qui contredit que ce soit un reste de division multivariée (Definition 4). □

Exercice 11 Montrer que la réciproque est vraie : soit $I = \langle G_1, \ldots, G_r \rangle$ un idéal. Si le reste de la division de tout polynôme $P \in I$ par $(G_1, \ldots, G_r)$ est nul, alors $(G_1, \ldots, G_r)$ est une base de Gröbner de I.

Théorème 6 Le reste R de la division d'un polynôme P par une base de Gröbner $(G_1, \ldots, G_r)$ d'un idéal I ne dépend que de I. On dit que R est *le* reste de la division de P par *l'idéal* I (relativement à l'ordre considéré). On a $R = 0$ si et seulement si $P \in I$.

Preuve. Soient R et R' deux restes obtenus dans la division de P en changeant l'ordre des G_i ou en considérant deux bases de Gröbner différentes de I. On a donc $R - R' \in I$. Si $R \neq R'$, on a donc $\mathrm{lt}(R - R') \in \mathrm{lt}(I)$, donc l'un des $\mathrm{lt}(G_i)$ divise $\mathrm{lt}(R - R')$ par propriété des bases de Gröbner. Or $\mathrm{lt}(R - R')$ provient nécessairement de l'un des termes de R ou R', disons R. Ceci implique R n'est pas une réduction modulo I, contradiction. □

Remarque 4 Si R est non nul, on a $\mathrm{lt}(R) < \mathrm{lt}(G)$ pour tout $G \in I$, ou de manière équivalente $\mathrm{lt}(R) < \mathrm{lt}(G_i)$ pour tout i. Visuellement, cela signifie que le support (ensemble des exposants) de G est strictement sous "l'escalier" déterminé par les exposants dominants des G_i.

Théorème 7 (Décomposition du quotient) Soit I un idéal de $\mathbb{K}[X_1, \ldots, X_n]$. On a un isomorphisme de $\mathbb{K}$-espaces vectoriels

$$\mathbb{K}[X_1, \ldots, X_n]/(I) \simeq \bigoplus_{M \in \mathcal{M}_n \setminus \mathrm{lm}(I)} \mathbb{K}\, M$$

Preuve. D'après la proposition précédente, chaque classe $P \mod I$ est uniquement représenté par le reste R de la division de P par I, et ce reste est une combinaison $\mathbb{K}$-linéaire de monômes $M \notin \mathrm{lm}(I)$. D'autre part, aucune combinaison linéaire non triviale de monômes $M \notin \mathrm{lm}(I)$ ne peut-être dans I. D'où l'isomorphisme. □

4.3 Existence des bases de Gröbner : l'algorithme de Buchberger

On établit de manière constructive l'existence des bases de Gröbner. On suit [BCGLLSS]. On se fixe un bon ordre monomial.

Définition 6 Soient $G, H \in \mathbb{K}[X_1, \ldots, X_n]$ non nuls. On appelle polynôme de syzygie ou S-polynôme de G et H le polynôme

$$S(G, H) = \frac{\mathrm{lt}(G)H - \mathrm{lt}(H)G}{\mathrm{pgcd}(\mathrm{lm}(G), \mathrm{lm}(H))} \in \mathbb{K}[X_1, \ldots, X_n]$$

On remarque que $S(G,H)$ est bien un polynôme, et que ce polynôme appartient à l'idéal engendré par G et H.

Par exemple, pour l'ordre lexico de $\mathbb{K}[X,Y]$ et $G = 2X^2Y - X$ et $H = 3X - Y^2$, on a

$$S(G,H) = \frac{2X^2YH - 3XG}{\text{pgcd}(X^2Y, X)} = 2XYH - 3G = 3X - 2XY^3.$$

On a le théorème fondamental suivant :

Théorème 8 (Caractérisation des bases de Gröbner) Une famille $(G_1, \ldots, G_r)$ constitue une base de Gröbner de l'idéal I qu'elle engendre si et seulement si le reste de la division de $S(G_i, G_j)$ par $G_1, \ldots, G_r$ est nul pour tout $i \neq j$.

Preuve. L'implication directe est immédiate au vue de la Proposition 7. On admet la réciproque. □

L'algorithme de Buchberger découle de ce théorème. Il s'écrit de la façon suivante.

Entrée : Un système de générateurs F de l'idéal I.
Sortie : Une base de Gröbner G de l'idéal I.

1. $G \leftarrow F$.
2. Répéter :
   a. $G_0 \leftarrow G$
   b. Pour toute paire $P \neq Q$ de G_0 faire :
      i. Diviser $S(P,Q)$ par G_0
      ii. Si le reste R est non nul, aors $G \leftarrow G \cup \{R\}$

jusqu'à $G = G_0$.

Théorème 9 L'algorithme termine et retourne G une base de Gröbner de I.

Preuve. Pour la terminaison, il suffit de considérer l'idéal monomial $\text{lm}(G) = (\text{lm}(G_1), \ldots, \text{lm}(G_r)$. À chaque tour de la boucle l'idéal croît au sens large. Il est constant à partir d'un certain rang par noethérianité (toute chaîne croissante d'idéaux de $\mathbb{K}[X_1, \ldots, X_n]$ est stationnaire). Or, si $\text{lm}(G)$ stagne, G stagne. En effet un reste R non nul agrandirait l'idéal monomial (son monôme de tête ne pourrait pas être dans $\text{lm}(G_0)$ car R est réduit modulo G_0 en fin de division). La correction de l'algorithme est un corollaire du théorème précédent : si l'algorithme s'arrête, c'est que tous les polynômes de syzygies se réduisent à zéro modulo G. □

5 Applications des Bases de Gröbner

On a vu que les bases de Gröbner permettent de trouver une base "canonique" (relativement à un bon ordre monomial) de l'algèbre quotient $R = \mathbb{K}[X_1, \ldots, X_n]/I$ et partant, d'effectuer des calculs algébriques modulo l'idéal. Ce résultat fondamental est à la base de l'algorithmique des systèmes polynomiaux (résolution, dimension, décomposition, singularités, etc). On énonce ici quelques applications importantes.

5.1 Variétés et idéaux

On appelle sous-variété algébrique (affine) $V \subset \mathbb{K}^n$ le lieu des zéros communs d'un système de polynômes $G_1, \ldots, G_r \in \mathbb{K}[X_1, \ldots, X_n]$ à plusieurs variables,

$$V = V(G_1, \ldots, G_r) := \{x \in \mathbb{K}^n,\ G_1(x) = \cdots = G_r(x) = 0\}.$$

Il est aisé de voir que tout polynôme dans l'idéal $I = \langle G_1, \ldots, G_r \rangle$ s'annule sur la variété V. On a donc l'égalité

$$V = V(I) := \{x \in \mathbb{K}^n,\ P(x) = 0, \forall\, P \in I\}.$$

Ainsi, l'idéal I est l'objet naturel attaché à la variété $G_1 = \cdots = G_r = 0$ et l'on parlera en général de la variété affine $V(I)$ définie par un idéal I. Cette variété est indépendante du choix des générateurs. Attention, la variété ainsi définie dépend par contre du choix du corps sur lequel on regarde les solutions.

Exemple 1. La sous-variété $V(X^2 + Y^2 + 1)$ de $\mathbb{R}^2$ est vide, mais elle ne l'est pas en tant que sous-variété de $\mathbb{C}^2$ (par exemple le point de coordonnées $(i, 0)$ est dessus).

Exemple 2. La variété $V(X^2 + Y^2 + Z^2 - 1, X - Y) \subset \mathbb{R}^3$ est une courbe dans l'espace (un cercle), intersection des deux surfaces $X^2 + Y^2 + Z^2 - 1 = 0$ (sphère unité) et $X - Y = 0$ (plan oblique).

Le Nullstellensatz (cf Section 5.6) assure que les relations idéaux et variétés sont dans un certain sens duales l'une de l'autre. Ce phénomène est à la base du principe fondamental de la géométrie algébrique :

Géométrie (étude des solutions des systèmes polynomiaux)
$$\longleftrightarrow$$
Algèbre (étude des idéaux de polynômes)

Les bases de Gröbner offrent un outil algorithmique performant pour établir cette passerelle. Illustrons ce fait avec une ou deux applications notables.

5.2 Caractérisation des systèmes avec un nombre fini de solutions

On suppose ici $\mathbb{K}$ algébriquement clos. On fixe un bon ordre monomial sur l'ensemble $\mathcal{M}_n$ des monômes de $\mathbb{K}[X_1, \ldots, X_n]$.

Proposition 8 Soit $I \subset \mathbb{K}[X_1, \ldots, X_n]$ un idéal. Les propositions suivantes sont équivalentes :

1. La variété $V(I)$ consiste en un nombre fini de points.
2. L'algèbre quotient $\mathbb{K}[X_1, \ldots, X_n]/(I)$ est de dimension finie sur $\mathbb{K}$.
3. L'ensemble $\mathcal{M}_n \setminus \mathrm{lm}(I)$ est fini.

On dit dans ce cas que I est un idéal zéro-dimensionnel.

Preuve. On a $2 \iff 3$ d'après le théorème 7. Si 2 est vraie, alors pour tout k, il existe $N \in \mathbb{N}$ tel que la famille $1, X_k, \cdots, X_k^N$ est liée dans l'algèbre quotient, i.e. il existe un polynôme $P_k \in \mathbb{K}[X_k]$ tel que $P_k \in I$. Il s'ensuit que $V(I) \subset V(P_1, \ldots, P_n)$. Or ce dernier ensemble est nécessairement fini (exo évident). Donc $2 \Rightarrow 1$. Réciproquement, si 1 est vrai, alors la projection de $V(I)$ sur l'axe des X_k est finie pour tout k. Il existe donc $P_k \in \mathbb{K}[X_k]$ qui s'annule sur cette projection. Par le Nullstellensatz (Theorem 14 ci-dessous), il existe donc $N_k \in \mathbb{N}$ tel que $Q_k = P_k^{N_k} \in I \cap \mathbb{K}[X_k]$. En particulier, $\langle Q_1, \ldots, Q_n \rangle \subset I$. Puisque l'algèbre quotient $\mathbb{K}[X_1, \ldots, X_n]/\langle Q_1, \ldots, Q_n \rangle$ est de dimension finie (exo), il en est donc de même pour $\mathbb{K}[X_1, \ldots, X_n]/(I)$. □

Proposition 9 Soit $I \subset \mathbb{K}[X_1, \ldots, X_n]$ un idéal. Les propositions suivantes sont équivalentes :

1. La variété $V(I)$ est vide.
2. L'idéal I est l'anneau $\mathbb{K}[X_1, \ldots, X_n]$ tout entier.
3. $1 \in \mathrm{lm}(I)$.

Preuve. La preuve est dans le même esprit (exo).

5.3 Elimination (ordre lexicographique)

Algébriquement, la « résolution » d'un système polynomial à plusieurs variables se ramène à une question d'élimination, déjà illustrée avec le résultant dans le cas de deux variables. Lorsqu'elle est possible, l'élimination successive des variables amène le système d'entrée sous une forme "triangulaire". De proche en proche, la résolution d'un tel système se réduit alors à la manipulation de polynômes à une variable, pour lesquels compter, isoler, voire calculer les solutions est bien plus facile.

Exemple. Considérons le système polynomial $F_1 = F_2 = F_3 = 0$ déterminé par les polynômes suivants

$$\begin{cases} F_1 = X^2 + Y + Z - 1 \\ F_2 = X + Y^2 + Z - 1 \\ F_3 = X + Y + Z^2 - 1 \end{cases}$$

Relativement à l'ordre lexicographique, l'idéal $I = \langle F_1, F_2, F_3 \rangle$ de $\mathbb{K}[X, Y, Z]$ admet pour base de Gröbner (G_1, G_2, G_3, G_4), avec

$$\begin{cases} G_1 = X + Y + Z^2 - 1 \\ G_2 = Y^2 - Y - Z^2 + Z \\ G_3 = 2YZ^2 + Z^4 - Z^2 \\ G_4 = Z^6 - 4Z^4 + 4Z^3 - Z^2 \end{cases}$$

Ainsi, les systèmes $F_1 = F_2 = F_3 = 0$ et $G_1 = \cdots = G_4 = 0$ ont même solutions (plus précisément, ils définissent la même variété). Puisque G_4 ne dépend que de la variable Z, on voit que les solutions ont nécessairement pour Z-coordonnées les racines de G_4. On substitue les racines z de G_4 dans G_2 et G_3, puis pour chaque z, on calcule les racines communes y des polynômes univariés $G_2(Y, z)$ et $G_3(Y, z)$ (i.e. racines du pgcd). On substitue alors chaque paire (y, z) dans G_1 et l'on calcule les racines x du polynôme univarié $G_1(X, y, z)$. Les solutions du système sont les triplets (x, y, z) obtenus.

Le point crucial ici est que la base de Gröbner a permis de "trianguler" notre système de départ. Le théorème fondamental suivant assure que ce fait est systématique dès lors que l'on choisit un ordre monomial adéquat, en l'occurence l'ordre lexico.

Théorème 10 (Théorème d'élimination) Soit $I \subset \mathbb{K}[X_1, \ldots, X_n]$ un idéal de base de Gröbner $\mathcal{G}$ pour l'ordre lexicographique. Le k-ème idéal éliminant de I

$$I_k := I \cap \mathbb{K}[X_{k+1}, \ldots, X_n]$$

est un idéal de $\mathbb{K}[X_{k+1}, \ldots, X_n]$ qui admet pour base de Gröbner $\mathcal{G}_k := \mathcal{G} \cap \mathbb{K}[X_{k+1}, \ldots, X_n]$.

Dans l'exemple précédent, ce résultat assure donc que l'on a les relations

$$I_1 = I \cap \mathbb{K}[Y, Z] = \langle G_2, G_3, G_4 \rangle \qquad \text{et} \qquad I_2 = I \cap \mathbb{K}[Z] = \langle G_4 \rangle$$

à partir desquelles on résoud le système "en remontant". Le théorème assure que ceci fonctionne en toute généralité. Noter la puissance des bases de Gröbner : non seulement les racines de G_4 donnent les z-coordonnées des solutions du système mais G_4 est le plus petit tel polynôme en tant que *générateur* de l'idéal I_2.

Remarque 5 Dans le cas $I = (A, B) \subset \mathbb{K}[X, Y]$ (deux équations à deux inconnues), le pgcd des éléments de $\mathcal{G} \cap \mathbb{K}[X]$ est donc un générateur de l'idéal principal $I \cap \mathbb{K}[X]$ (on considère ici l'ordre lexicographique inversé car on élimine Y). Il donne donc une information plus fine que le résultant $\mathrm{Res}_Y(A, B)$ qui certes vit dans l'idéal $I \cap \mathbb{K}[X]$, mais n'en est pas nécessairement un générateur.

5.4 Relèvement des solutions.

Comme pour le résultant, il se peut qu'étant donnés des points

$$x'' = (x_{k+1}, \ldots, x_n)$$

dans la sous-variété $V(I_k) \subset \mathbb{K}^{n-k}$, il n'existe pas de solutions du système $x \in V(I) \subset \mathbb{K}^n$ de la forme $x = (x', x'')$, i.e. qui se projettent sur x'' via la projection

$$\pi : \mathbb{K}^n \simeq \mathbb{K}^k \times \mathbb{K}^{n-k} \to \mathbb{K}^{n-k}.$$

Autrement dit, la projection $\pi : V(I) \to V(I_k)$ n'est pas nécessairement surjective. Ce phénomène s'explique par le fait que $V(I_k)$ capture également les "solutions à l'infini" (x', x'') du système pour lesquelles certaines des coordonnées de x' sont infinies.

Exemple. Soit $I = \langle X_1X_2 - 1, X_2\rangle$. On a $I_1 = I \cap \mathbb{K}[X_2] = \langle X_2\rangle$ et $V(I_2)$ est réduite au point $x_2 = 0$. En remplaçant X_2 par 0 dans la première équation, on obtient $-1 = 0$ qui n'a pas de solutions. Donc $V(I)$ est vide, le système n'a pas de solutions. Le fait que $V(I_1)$ contienne malgré tout une solution $x_2 = 0$ s'interprète ici par le fait que l'hyperbole $X_1X_2 - 1$ et la droite horizontale $X_2 = 0$ ont une même asymptote horizontale. Autrement dit, elles "s'intersectent à l'infini" en le point $(x_1, x_2) = (\infty, 0)$ qui "se projette" sur $x_2 = 0$.

Le résultat suivant permet de caractériser les "mauvaises valeurs" de la variété $V(I_1)$ du premier idéal éliminant. Observer l'analogie avec le résultant, qui d'ailleurs apparaît dans la preuve (admise).

Théorème 11 (Théorème de relèvement) Suppose $\mathbb{K}$ algébriquement clos. Soit $I = \langle F_1, \ldots, F_s\rangle$ un idéal de $\mathbb{K}[X_1, \ldots, X_n]$ et soit $I_1 = I \cap \mathbb{K}[X_2, \ldots, X_n]$ son premier idéal éliminant. Pour tout point $(x_2, \ldots, x_n) \in V(I_1)$ qui n'annule pas tous les termes de tête des F_i (vus comme polynômes en x_1), il existe une solution $x \in V(I)$ de la forme $x = (x_1, x_2, \ldots, x_n)$.

Puisque I_r est le premier idéal éliminant de I_{r+1}, ce résultat suffit pour éliminer les "mauvaises valeurs" pas à pas en considérant de droite à gauche les projections successives

$$V(I) \to V(I_1) \to V(I_2) \to \cdots \to V(I_{n-1}),$$

et en regardant pour chaque $k = n, \ldots, 1$ les coefficients de tête (pour la variable X_k) des générateurs de l'idéal éliminant I_{k-1}.

Exercice 12 Calculer les idéaux d'éliminations de l'idéal

$$I = \langle X^2 + Y^2 + Z^2 - 4, X^2 + 2Y^2 - 5, XZ - 1\rangle.$$

Résoudre le système d'équations associé dans $\mathbb{Q}^3$ puis dans $\mathbb{R}^3$.

5.5 Implicitation des variétés paramétrées

On considère une paramétrisation polynomiale $\Phi : \mathbb{K}^m \to \mathbb{K}^n$ définie par

$$\Phi : (t_1, \ldots, t_m) \longmapsto (P_1(t_1, \ldots, t_m), \ldots, P_n(t_1, \ldots, t_m))$$

où $P_i \in \mathbb{K}[T_1, \ldots, T_m]$. On s'intéresse à la plus petite sous-variété $V \subset \mathbb{K}^n$ contenant l'image $W = \Phi(\mathbb{K}^m)$ de la paramétrisation. On appelle cette variété V la *clôture de Zariski* de W.

Remarque 6 Il peut arriver que l'inclusion $W \subset V$ soit stricte : penser à $\Phi : \mathbb{R} \to \mathbb{R}$ donnée par $t \mapsto t^2$. Dans ce cas, $W = \mathbb{R}^+$ n'est pas une sous-variété de $\mathbb{R}$: il n'existe pas de famille de polynômes de $\mathbb{R}[X]$ dont le lieu des zéros est cette demi-droite. Il est facile de voir ici que la clôture de Zariski de $\mathbb{R}^+$ est la droite affine $\mathbb{R}$.

Le *problème d'implicitation* consiste à déterminer un idéal I tel que $V(I) = V$. Autrement dit, déterminer une famille (finie) de polynômes qui s'annulent sur W et telle que tout autre polynôme s'annulant sur W est dans l'idéal engendré par cette famille.

De manière analogue au cas des courbes planes paramétrées (cf Chapitre sur les résultants), on résoud ce problème en introduisant des indéterminées $X_1, \ldots, X_n$ et en considérant les polynômes $F_i \in \mathbb{K}[T_1, \ldots, T_m, X_1, \ldots, X_n]$ définis par

$$\begin{cases} F_1 = X_1 - P_1(T_1, \ldots, T_m) \\ F_2 = X_2 - P_1(T_1, \ldots, T_m) \\ \vdots \\ F_n = X_n - P_n(T_1, \ldots, T_m) \end{cases}$$

Le théorème suivant permet de ramener le problème d'implicitation à un problème d'élimination.

Théorème 12 (Implicitation polynomiale) . Soit $\mathbb{K}$ un corps infini et soit $\Phi : \mathbb{K}^m \to \mathbb{K}^n$ la paramétrisation polynomiale ci-dessus. Soit $J = \langle F_1, \ldots, F_n \rangle \subset \mathbb{K}[T_1, \ldots, T_m, X_1, \ldots, X_n]$ l'ideal engendré par les F_i ci-dessus et soit

$$I = J \cap \mathbb{K}[X_1, \ldots, X_n]$$

le m-ème idéal éliminant de I. Alors la sous-variété $V(I) \subset \mathbb{K}^n$ est la plus petite sous-variété contenant l'image $W = \Phi(\mathbb{K}^m)$.

Combiné avec le théorème d'élimination, ce théorème assure alors que les bases de Gröbner relativement à l'ordre *lexicographique* $T_1 \succ \cdots \succ T_m \succ X_1 \succ \cdots \succ X_n$ résolvent efficacement notre problème d'implicitation :

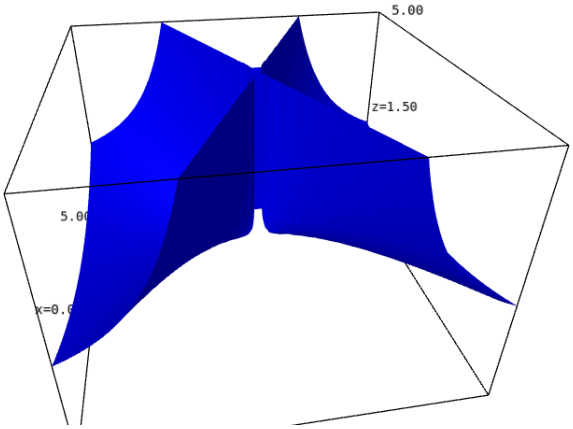
1. Calculer une base de Gröbner $\mathcal{G}$ de l'idéal J
2. Retourner l'idéal I engendré par les éléments de $\mathcal{G} \cap \mathbb{K}[X_1, \ldots, X_n]$.

Pour vérifier si Φ paramétrise bien la variété V tout entière (paramétrisation surjective d'une variété), i.e. si $W = V$, il suffit d'utiliser le théorème d'extension afin de vérifier que toute solution $x \in V(I)$ s'étend bien en une solution $(t, x) \in V(J)$: ceci équivaut au fait que x est dans l'image de Φ par définition de J.

Exercice 13 Calculer à la main puis avec les bases de Gröbner l'équation implicite de la surface algébrique paramétrée par

$$\Phi : (u, v) \longmapsto (uv, v, u^2).$$

Montrer que la paramétrisation couvre toute la surface si $\mathbb{K} = \mathbb{C}$. Montrer qu'il manque des points si $\mathbb{K} = \mathbb{R}$ (et déterminer les points manquants). Cette surface s'appelle le parapluie de Whitney, représenté ci-dessous.



Remarque 7 (Implicitation rationnelle) Le problème d'implicitation s'étend au cadre des paramétrisations rationnelles

$$\Phi : t = (t_1, \ldots, t_m) \longmapsto \left(\frac{P_1(t)}{Q_1(t)}, \ldots, \frac{P_n(t)}{Q_n(t)} \right)$$

où les P_i et Q_i sont des polynômes de $\mathbb{K}[T_1, \ldots, T_m]$. On considère dans ce cas l'image de Φ restreinte à son domaine de définition $Q_1 \cdots Q_n \neq 0$. Pour trouver la clôture de Zariski de cette image, on raisonne de manière similaire au cas polynomial, mais en ajoutant astucieusement une nouvelle indéterminée et une nouvelle équation : on considère cette fois l'idéal

$$J = \langle Q_1 X_1 - P_1, \ldots, Q_n X_n - P_n, Q_1 \cdots Q_n Y - 1 \rangle.$$

On peut montrer que la plus petite variété contenant l'image de Φ est définie par l'idéal éliminant $J = I \cap \mathbb{K}[X_1, \ldots, X_n]$. A nouveau, les bases de Gröbner permettent de calculer cet idéal efficacement.

5.6 Nullstellensatz (appendice)

On a vu comment attacher à tout idéal I un sous-ensemble $V(I)$ de $\mathbb{K}^n$. Réciproquement, on peut attacher à un sous-ensemble $Z \subset \mathbb{K}^n$ son idéal annulateur

$$\mathbf{I}(Z) = \{G \in \mathbb{K}[X_1, \ldots, X_n],\ G(x) = 0\ \forall x \in Z\}.$$

Lemme 2 On a les propriétés suivantes :

1. Pour deux idéaux $I \subset J$, on a $V(I) \supset V(J)$
2. Pour deux sous-ensembles $W \subset Z$, on a $\mathbf{I}(W) \supset \mathbf{I}(Z)$
3. Pour tout idéal I, on a $I \subset \mathbf{I}(V(I))$
4. Pour tout sous-ensemble Z on a $Z \subset V(\mathbf{I}(Z))$.

Exercice 14 Prouver ce lemme.

Malheureusement, les opérateurs V et $\mathbf{I}$ ne sont pas exactement "duaux" l'un de l'autre, dans le sens ou l'inclusion du point 3 du lemme n'est en général pas une égalité. Par exemple, si $I = \langle X^2 + Y, Y \rangle$, on a $V(I) = \{(0,0)\}$ de sorte que $X \in \mathbf{I}(V(I))$. Pour autant, $X \notin I$.

Le Nullstellensatz (en allemand, « théorème du lieu des zéros ») est le théorème qui complète le lemme précédent et établit un lien "presque dual" entre les variétés algébriques (affines) et les idéaux. Il en découle que certaines opérations géométriques se traduisent aisément en algorithmes sur les équations. Par exemple, le Nullstellensatz est utile à la preuve de la caractérisation des idéaux de dimension zéro donnée précédemment.

Définition 7 Le radical d'un idéal $I \subset \mathbb{K}[X_1, \ldots, X_n]$ est défini par

$$\sqrt{I} = \{f \in \mathbb{K}[X_1, \ldots, X_n],\ \exists\, k \in \mathbb{N},\ f^k \in I\}.$$

C'est un idéal qui contient I. On dit que I est radical si $I = \sqrt{I}$.

Par exemple, l'idéal $I = \langle X^2 \rangle \subset \mathbb{K}[X]$ admet pour radical $\sqrt{I} = \langle X \rangle$.

On admettra les deux théorèmes fondamentaux suivants, dont une preuve repose sur la théorie des résultants (tout à fait accessible, cf Section 25 de [BCLLSS] par exemple).

Théorème 13 (Nullstellensatz faible) Si $\mathbb{K}$ est algébriquement clos, alors $V(I) = \emptyset$ si et seulement si $I = \mathbb{K}[X_1, \ldots, X_n]$, i.e. si et seulement si $1 \in I$.

Un sens est évident : si le polynôme constant $P = 1$ appartient à I, il est clair que $V(I)$ est vide. L'autre sens est plus délicat. Il assure que si $G_1, \ldots, G_r$ n'ont aucun zéros en commun, il y a nécessairement une relation "de type Bezout"

$$U_1 G_1 + \cdots + U_r G_r = 1$$

dans l'anneau de polynôme $\mathbb{K}[X_1, \ldots, X_n]$ (généralisation du théorème de Bezout dans le cadre de plusieurs variables).

Théorème 14 (Nullstellensatz) Si $\mathbb{K}$ est algébriquement clos, on a l'égalité $\mathbf{I}(V(I)) = \sqrt{I}$.

Autrement dit, si G s'annule identiquement sur $V(I)$, alors on est sûr qu'une puissance de G appartient à I. Par exemple, le polynôme X s'annule sur $V(X^2 + Y, Y)$ et, bien que X ne soit pas dans l'idéal, il est dans son radical puisque $X^2 \in \langle X^2 + Y, Y \rangle$.

Voilà pour quelques premiers pas dans le vaste monde de la géométrie algébrique...