Élimination, résultant. Discriminant

Michel Coste *

Version révisée le 9 Octobre 2001

Introduction

Le programme de l'agrégation (pour la session 2002) mentionne « Résultant et discriminant » dans le paragraphe commençant par « Relations entre les coefficients et les racines... ». Le mot « Résultant » ne figure plus dans les intitulés de leçons d'algèbre et géométrie (mais on y retrouve toujours les relations entre les coefficients et les racines d'un polynôme). Par ailleurs, le programme de l'option Calcul numérique et symbolique de l'épreuve de modélisation mentionne le « calcul effectif des résultants, application à l'élimination », et un intitulé de leçon est « Résultant et élimination effective dans les systèmes d'équations polynomiales. Application(s) issue(s) des thèmes du programme. ».

Les ouvrages classiques de préparation à l'agrégation qui traitent le plus longuement de résultant et d'élimination sont les anciens manuels de classes préparatoires (par exemple [LeAr], chapitre 6, [Qu], chapitre 13). Le sujet de l'élimination apparaît un peu vieillot, marqué « 19ème siècle ». Il ne figure ni dans l'ouvrage de Perrin, ni dans le Algebra de M. Artin. Il est traité, assez rapidement, chez Lang [La], chapitre 5, et Tauvel [Ta], chapitre 13. Résultant et discriminant sont traités à la fin du chapitre IV du traité d'Algèbre de Bourbaki [Bo] (pas dans les anciennes éditions). Le sujet de l'élimination a repris récemment un peu de vigueur avec le développement du calcul formel ([Mi] chapitre 3, [Sa] chapitre 7). On n'a retenu dans les références que les ouvrages contenant au moins les points fondamentaux : présentation du résultant comme déterminant de la matrice de Sylvester, théorème principal (théorème 3 dans le texte qui suit), expression en fonction des racines.

On peut trouver dans les anciens manuels de classes préparatoires une définition de l'élimination. Voici par exemple celle qui figure dans [Qu] :

> Soit deux équations algébriques
>
> $$\begin{aligned} P(X) = a_0X^p + a_1X^{p-1} + \cdots + a_0 &= 0 \qquad (a_0 \neq 0) \\ Q(X) = b_0X^q + b_1X^{q-1} + \cdots + b_q &= 0 \qquad (b_0 \neq 0). \end{aligned}$$
>
> Éliminer X entre ces deux équations, c'est trouver une condition vérifiée par les coefficients des deux équations, nécessaire et suffisante pour que ces deux équations aient au moins une racine commune.

On peut se poser la question : où doit être la racine commune ? Si l'on suppose que P et Q sont à coefficients dans un corps K, on pourrait demander que la racine soit dans ce même corps K. Mais dans ce cas on ne sait pas faire grand chose en général. Par contre, le problème est plus commode si on demande que P et Q aient une racine commune dans un corps algébriquement clos $\overline{K}$ contenant K. Dans ce cas, on a une réponse facile.

Proposition 1 *Les polynômes P et Q ont une racine commune dans $\overline{K}$ si et seulement s'ils ont un diviseur commun non constant dans $K[X]$, c'est à dire si et seulement si leur pgcd est de degré ≥ 1.*

Démonstration : Ceci vient simplement du fait que P et Q ont une racine commune dans $\overline{K}$ si et seulement si leur pgcd dans $\overline{K}[X]$ est de degré ≥ 1, et du fait qu'un pgcd D de P et Q dans $K[X]$ est aussi pgcd dans $\overline{K}[X]$ (il divise P et Q, et il existe U et V tels que $D = UP + VQ$). □

*Merci à L. Moret-Bailly et J-C. Raoult pour leurs remarques.

On pourrait penser que le problème de l'élimination est résolu, puisqu'on sait décider si deux polynômes donnés ont ou non une racine commune. Mais visiblement, ce n'est pas ce qu'on cherche : on a en vue des polynômes avec des coefficients dépendant de paramètres (par exemple, des polynômes à plusieurs variables, où on privilégie une variable), et alors le calcul du pgcd ne peut en général pas être fait d'une manière uniforme pour tous les paramètres. La réponse au problème est alors fournie par le *résultant* de P et de Q, qui est un polynôme à coefficients entiers en les coefficients $(a_0, \ldots, a_p, b_0, \ldots, b_q)$, et qui s'annule si et seulement si P et Q ont une racine commune, ou si a_0 et b_0 sont simultanément nuls.

On peut faire une analogie avec la situation que l'on connait pour les systèmes de n équations linéaires à n inconnues. Le moyen le plus commode de résoudre un tel système est le pivot de Gauss. Mais si l'on discute un système où les coefficients dépendent de paramètres, il est utile de faire intervenir le déterminant du système. Cette analogie n'est pas superficielle ; on pourrait approfondir les rapports entre l'algorithme d'Euclide de calcul du pgcd et le pivot de Gauss (voir l'exercice 3), et on verra que le résultant est effectivement défini comme un déterminant.

L'étude du discriminant se rattache naturellement à celle du résultant. Le problème ici est de trouver une condition nécessaire et suffisante sur les coefficients du polynôme P pour que celui-ci ait une racine multiple (dans un corps algébriquement clos $\overline{K}$ contenant K). Ceci revient à demander que P et P' aient un facteur commun non constant. On pourrait définir le discriminant de P comme le résultant de P et P', mais la tradition (pas toujours respectée dans les ouvrages) donne une autre définition. Au-delà de ces divergences de définition, le point à retenir est qu'un polynôme unitaire P a ses racines distinctes si et seulement si son discriminant est non nul, et que ce discriminant est un polynôme à coefficients entiers en les coefficients de P.

1 Résultant : définition et propriétés fondamentales

On a dit dans l'introduction que les techniques d'élimination étaient surtout intéressantes dans le cas où les coefficients des polynômes dépendent de paramètres. Ceci amène à considérer des polynômes $P = u_0X^p + \cdots + u_p$ et $Q = v_0X^q + \cdots + v_q$ à coefficients indéterminés, de degrés $p > 0$ et $q > 0$ respectivement. Notons $\underline{u} = (u_0, \ldots, u_p)$ et $\underline{v} = (v_0, \ldots, v_q)$. Les polynômes P et Q appartiennent à $\mathbb{Z}[\underline{u}, \underline{v}, X]$.

Soit A un anneau commutatif unitaire. On peut substituer à $\underline{u}$ et $\underline{v}$ des suites $\underline{a} = (a_0, \ldots, a_p)$ et $\underline{b} = (b_0, \ldots, b_q)$ d'éléments de A, et on obtient ainsi des polynômes $P_{\underline{a}}$ et $Q_{\underline{b}}$ de $A[X]$; on dit qu'on a spécialisé $\underline{u}$ et $\underline{v}$ en $\underline{a}$ et $\underline{b}$.

Le *résultant* de P et Q est le déterminant de la *matrice de Sylvester* de P et Q, qui est la matrice carrée de taille $p + q$:

$$\begin{pmatrix}
u_0 & u_1 & \ldots & \ldots & \ldots & u_p & 0 & \ldots & 0 \\
0 & u_0 & u_1 & \ldots & \ldots & \ldots & u_p & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \ldots & 0 & u_0 & u_1 & \ldots & \ldots & \ldots & u_p \\
v_0 & v_1 & \ldots & \ldots & v_q & 0 & \ldots & \ldots & 0 \\
0 & v_0 & v_1 & \ldots & \ldots & v_q & 0 & \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \ldots & \ldots & 0 & v_0 & v_1 & \ldots & \ldots & v_q
\end{pmatrix}$$

Les lignes de la matrice de Sylvester sont les coefficients des polynômes $X^{q-1}P$, $X^{q-2}P$,..., P, $X^{p-1}Q$,..., Q rapportés à X^{p+q-1}, ..., X, 1. On notera $\mathrm{res}(P, Q)$ le résultant de P et de Q. S'il est utile de préciser le nom de l'indéterminée, on notera $\mathrm{res}_X(P, Q)$. Ce résultant est un polynôme, disons $R(\underline{u}, \underline{v})$ de $\mathbb{Z}[\underline{u}, \underline{v}]$. Si maintenant on spécialise $\underline{u}$ et $\underline{v}$ en des suites $\underline{a}$ et $\underline{b}$ d'éléments de A, on obtient un élément $R(\underline{a}, \underline{b})$ de A, que l'on désignera comme le résultant de $P_{\underline{a}}$ et $Q_{\underline{b}}$, et que l'on notera $\mathrm{res}(P_{\underline{a}}, Q_{\underline{b}})$. Ce résultant se calcule comme le déterminant de la matrice ci-dessus, où les $\underline{u}$ et $\underline{v}$ sont remplacés par les $\underline{a}$ et $\underline{b}$.

Exercice 1 Montrer que $\mathrm{res}(Q, P) = (-1)^{pq}\mathrm{res}(P, Q)$.

Exercice 2 Comparer le polynôme obtenu en faisant $v_0 = 0$ dans le résultant de $P = u_0X^p + \cdots + u_p$ et $Q = v_0X^q + \cdots + v_q$, et le résultant des polynômes P et $v_1X^{q-1} + \cdots + v_q$. (Le premier vaut u_0 fois le deuxième.)

Ce dernier exercice met le doigt sur une difficulté dans la définition du résultant : le résultat n'est pas exactement le même suivant que l'on considère Q comme polynôme de degré q où l'on spécialise le coefficient dominant à 0, ou comme polynôme de degré $q-1$. Dans le cas de polynômes à coefficients dépendant de paramètres, il se peut que le degré baisse pour certaines valeurs des paramètres, mais il est important que la formule qui calcule le résultant soit la même pour tous les paramètres. Ceci est un argument en faveur de la présentation choisie (considérer des polynômes à coefficients indéterminés, que l'on spécialise ensuite ; c'est l'approche de [La]), par rapport à l'approche a priori plus simple qui consisterait à définir directement le résultant de deux polynômes à coefficients « fixés » (comme par exemple dans [Ta]). Dans [Bo] on utilise la notation $\mathrm{res}_{p,q}(f,g)$ pour indiquer qu'on calcule le résultant de f et g comme spécialisations de polynômes généraux de degrés p et q, même si les degrés réels de f et g sont plus petits.

Commençons par une propriété importante du résultant.

Proposition 2 *Il existe des polynômes $\Lambda(\underline{u},\underline{v},X)$ de degré $< q$ en X et $\Theta(\underline{u},\underline{v},X)$ de degré $< p$ en X dans $\mathbb{Z}[\underline{u},\underline{v},X]$ tels que $\mathrm{res}(P,Q) = \Lambda P + \Theta Q$. En particulier, $\mathrm{res}(P,Q)$ appartient à l'idéal engendré par P et Q dans $\mathbb{Z}[\underline{u},\underline{v},X]$*

Démonstration : On sait qu'on ne change pas la valeur d'un déterminant en ajoutant à une colonne une combinaison linéaire des autres colonnes. Ici, on considère le déterminant $\mathrm{res}(P,Q)$ dans l'anneau $\mathbb{Z}[\underline{u},\underline{v},X]$, et on ajoute à la dernière colonne X fois l'avant dernière plus X^2 fois l'antépénultième plus ... plus X^{p+q-1} fois la première. La dernière colonne devient alors

$$\begin{pmatrix} X^{q-1}P \\ \vdots \\ P \\ X^{p-1}Q \\ \vdots \\ Q \end{pmatrix},$$

et en développant le déterminant par rapport à cette colonne on obtient bien l'expression annoncée pour $\mathrm{res}(P,Q)$. □

Voici maintenant le théorème central de la théorie du résultant.

Théorème 3 *Soient*

$$\begin{aligned} P_{\underline{a}} &= a_0X^p + a_1X^{p-1} + \cdots + a_p, \\ Q_{\underline{b}} &= b_0X^q + b_1X^{q-1} + \cdots + b_q \end{aligned}$$

deux polynômes à coefficients dans un corps K. Soit $\overline{K}$ un corps algébriquement clos contenant K. Les propriétés suivantes sont équivalentes.

1. *Les polynômes $P_{\underline{a}}$ et $Q_{\underline{b}}$ ont une racine commune dans $\overline{K}$, ou $a_0 = b_0 = 0$.*
2. *Les polynômes $P_{\underline{a}}$ et $Q_{\underline{b}}$ ont un diviseur commun non constant dans $K[X]$, ou $a_0 = b_0 = 0$.*
3. *Il existe des polynômes à coefficients dans K, U de degré $< q$ et V de degré $< p$, non nuls tous les deux, tels que $UP_{\underline{a}} + VQ_{\underline{b}} = 0$.*
4. *$\mathrm{res}(P_{\underline{a}}, Q_{\underline{b}}) = 0$.*

Démonstration : L'équivalence de 1) et 2) a été vue.

Si $a_0 = b_0 = 0$ alors $\deg P_{\underline{a}} < p$ et $\deg Q_{\underline{b}} < q$, et l'égalité $Q_{\underline{b}}P_{\underline{a}} - P_{\underline{a}}Q_{\underline{b}} = 0$ montre 3). On suppose dans la suite que a_0 et b_0 ne sont pas simultanément nuls, par exemple $a_0 \neq 0$ (le raisonnement avec $b_0 \neq 0$ est symétrique).

Montrons que 2) entraîne 3). Soit D un diviseur commun non constant de $P_{\underline{a}}$ et $Q_{\underline{b}}$. En posant $U = Q_{\underline{b}}/D$ et $V = -P_{\underline{a}}/D$, on a bien la propriété 3).

Montrons que 3) entraîne 2). De $UP_{\underline{a}}+VQ_{\underline{b}} = 0$, on déduit que $P_{\underline{a}}$ divise VQ. Si $P_{\underline{a}}$ et $Q_{\underline{b}}$ sont premiers entre eux, alors $P_{\underline{a}}$ divise V, ce qui est impossible parce que V est non nul et que $\deg V < p = \deg P_{\underline{a}}$. Donc $P_{\underline{a}}$ et $Q_{\underline{b}}$ ne sont pas premiers entre eux, c.-à-d. qu'ils ont un diviseur commun non constant.

Montrons que 3) est équivalent à 4). La propriété 3) est vérifiée si et seulement s'il existe $\lambda_{q-1}, \ldots, \lambda_0, \mu_{p-1}, \ldots, \mu_0$ dans K, non tous nuls, tels que

$$\lambda_{q-1}X^{q-1}P_{\underline{a}} + \lambda_{q-2}X^{q-2}P_{\underline{a}} + \cdots + \lambda_0 P_{\underline{a}} + \mu_{p-1}X^{p-1}Q_{\underline{b}} + \mu_{p-2}X^{p-2}Q_{\underline{b}} + \cdots + \mu_0 Q_{\underline{b}} = 0.$$

On rapporte l'espace vectoriel des polynômes de $K[X]$ de degré $< p+q$ à la base $X^{p+q-1}, X^{p+q-2}, \ldots, X, 1$. Alors l'équation ci-dessus s'écrit comme un système de $p+q$ équations linéaires sans second membre en les $p+q$ inconnues λ_i et μ_j, et la matrice de ce système est la transposée de la matrice de Sylvester de $P_{\underline{a}}$ et $Q_{\underline{b}}$; le déterminant du système est donc $\mathrm{res}(P_{\underline{a}}, Q_{\underline{b}})$. La propriété 3) est vérifiée si et seulement si ce système admet une autre solution que la solution $(0, 0, \ldots, 0)$, c'est à dire si et seulement si $\mathrm{res}(P_{\underline{a}}, Q_{\underline{b}}) = 0$. □

La démonstration ci-dessus est la démonstration classique utilisant l'algèbre linéaire. On peut aussi utiliser la proposition 2. L'implication 4) ⇒ 3) est une conséquence immédiate de cette proposition (on trouve U et V en spécialisant $(\underline{u}, \underline{v})$ en $(\underline{a}, \underline{b})$ dans Λ et Θ).

L'énoncé du théorème fondamental dans [Ta] ne mentionne pas l'alternative $a_0 = b_0 = 0$ dans 2). Ceci vient du fait que la définition du résultant dans [Ta] suppose $a_0 \neq 0$ et $b_0 \neq 0$.

Le cas $a_0 = b_0 = 0$ peut d'ailleurs se voir comme l'existence d'une racine commune « à l'infini » pour les polynômes $P_{\underline{a}}$ et $Q_{\underline{b}}$. De manière précise, on homogénéise le polynôme $P_{\underline{a}}$ en posant $(P_{\underline{a}})^h(X, Y) = Y^p P_{\underline{a}}(X/Y)$, et on dit qu'un élément $(\alpha{:}\beta)$ de la droite projective $\mathbb{P}^1(\overline{K})$ est un zéro de $(P_{\underline{a}})^h$ si $(P_{\underline{a}})^h(\alpha, \beta) = 0$ (ceci ne dépend pas du représentant choisi d'après l'homogénéité). Les zéros $(\alpha{:}1)$ correspondent aux racines α de $P_{\underline{a}}$ dans $\overline{K}$, et le point à l'infini (1:0) est zéro de $(P_{\underline{a}})^h$ si et seulement si $a_0 = 0$.

Plusieurs ouvrages ([Mi], [Sa]) expliquent comment calculer le résultant par l'algorithme d'Euclide (en faisant des divisions euclidiennes successives). Ceci vaut pour deux polynômes à coefficients dans un corps. L'exercice suivant permet de voir comment ceci marche.

Exercice 3 Soient

$$\begin{aligned} P_{\underline{a}} &= a_0X^p + a_1X^{p-1} + \cdots + a_p, \\ Q_{\underline{b}} &= b_0X^q + b_1X^{q-1} + \cdots + b_q \end{aligned}$$

deux polynômes à coefficients dans un corps K, avec $a_0 \neq 0$. On fait la division euclidienne $Q_{\underline{b}} = P_{\underline{a}}F + R$, avec $\deg(R) = d < p$. Montrer que

$$\mathrm{res}(P_{\underline{a}}, Q_{\underline{b}}) = a_0^{q-d}\mathrm{res}(P_{\underline{a}}, R)\ .$$

Si c est une constante dans K, que vaut $\mathrm{res}(P_{\underline{a}}, c)$? Expliquer comment calculer le résultant en utilisant l'algorithme d'Euclide (utiliser aussi l'exercice 1)

On voit bien que cette méthode des divisions successives pose des problèmes de spécialisation quand les coefficients dépendent de paramètres. Il y a cependant une notion de « polynôme sous-résultant », liée à l'utilisation de « pseudo-divisions » dans l'algorithme d'Euclide, qui permet de contourner cette difficulté et qui est utilisée en pratique pour calculer les résultants (on lit dans la documentation de Maple que « the subresultant algorithm is used for polynomials of low degree »). Nous n'entrerons pas dans cette théorie des sous-résultants. Nous nous contentons dans l'exercice suivant de donner une propriété des coefficients dominants des polynômes sous-résultants (ce sont les $\mathrm{sr}_k(P, Q)$).

Exercice 4 On note $\mathrm{sr}_k(P, Q)$ le déterminant de la matrice carrée de taille $p + q - 2k$ extraite de la matrice de Sylvester de P et Q en supprimant les k dernières lignes de coefficients de P, les k dernières lignes de coefficients de Q, et les $2k$ dernières colonnes, ceci pour $0 \leq k < \inf(p, q)$. On a en particulier $\mathrm{sr}_0(P, Q) = \mathrm{res}(P, Q)$. Montrer que $\mathrm{sr}_0(P_{\underline{a}}, Q_{\underline{b}}) = \mathrm{sr}_1(P_{\underline{a}}, Q_{\underline{b}}) = \ldots = \mathrm{sr}_k(P_{\underline{a}}, Q_{\underline{b}}) = 0$ si et seulement si le degré du plus grand commun diviseur de $P_{\underline{a}}$ et $Q_{\underline{b}}$ est $> k$, ou $a_0 = b_0 = 0$.

On termine cette section par un aspect « géométrie différentielle » du résultant.

Exercice 5 On identifie l'ensemble des polynômes unitaires de degré n à coefficients dans $\mathbb{R}$ à l'espace $\mathbb{R}^n$ au moyen de la bijection $X^n+a_1X^{n-1}+\cdots+a_n \mapsto (a_1,\ldots,a_n)$. Moyennant cette identification, on note $\mu : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^{p+q}$ l'application qui aux polynômes unitaires A et B de degrés respectivement p et q fait correspondre leur produit AB. Comparer le déterminant jacobien de μ au point (A,B) et le résultant $\mathrm{res}(A,B)$. En déduire le résultat suivant : Si A et B sont premiers entre eux, il existe des voisinages U de A dans $\mathbb{R}^p$, V de B dans $\mathbb{R}^q$ et W de AB dans $\mathbb{R}^{p+q}$ tels que μ induise un difféomorphisme de $U \times V$ sur W. En particulier, pour tout F dans W il existe un unique couple (G,H) dans $U \times V$ tel que $F = GH$.

2 Applications du résultant

On présente sous forme d'exercices quelques exemples d'application du résultant. Ce sont les vieux ouvrages, notamment de classes préparatoires, qui sont les plus diserts sur ce sujet.

On peut voir dans l'exercice suivant comment le résultant peut être utilisé pour ramener un système de deux équations polynomiales en deux variables à des équations en une variable.

Exercice 6 Calculer le résultant $\mathrm{res}_Y(P,Q)$ des polynômes $P = X^2-XY+Y^2-1$ et $Q = 2X^2+Y^2-Y-2$ par rapport à la variable Y. Utiliser le résultat pour trouver les points d'intersection des ellipses d'équations $P = 0$ et $Q = 0$.

Il faut se souvenir, quand on applique le résultant à la résolution de systèmes polynomiaux comme dans l'exercice ci-dessus, de l'alternative $a_0 = b_0 = 0$ dans la propriété 1) du théorème 3. Par exemple, $\mathrm{res}_Y(XY-1, XY) = X$, mais la racine 0 du résultant ne se relève sûrement pas en une solution $(0,y)$ du système $XY - 1 = XY = 0$.

Le résultant peut servir à former un polynôme dont les racines sont des expressions algébriques d'une racine d'un polynôme P et d'une racine d'un polynôme Q.

Exercice 7 Soient A et B deux polynômes de $K[X]$, où K est un corps. Fabriquer un polynôme dont les racines sont les sommes d'une racine de A et d'une racine de B. (Quels sont les Y tels que le système $A(X) = B(Y-X) = 0$ ait une solution ?)

Fabriquer un polynôme à coefficients entiers qui a $\sqrt{2}+\sqrt[3]{7}$ pour racine.

Le résultant peut aussi servir pour ce qu'on appelle la « transformation » des équations (voir [LeAr]).

Exercice 8 . Soient A et P des polynômes de $K[X]$. Fabriquer en utilisant le résultant un polynôme dont les racines sont les $P(\alpha)$, pour α racine de A.

Le résultant peut servir à passer d'une paramétrisation rationnelle d'une courbe à son équation algébrique (implicitation).

Exercice 9 Comment fabriquer l'équation de la courbe paramétrée par $x = A(t)/B(t)$, $y = F(t)/G(t)$, où A, B, F, G sont des polynômes ? Exemple : $x = t^2+t+1$, $y = (t^2-1)/(t^2+1)$.

3 Expression du résultant en fonction des racines

On considère ici les racines $\underline{\alpha} = (\alpha_1,\ldots,\alpha_p)$ et $\underline{\beta} = (\beta_1,\ldots,\beta_q)$ des polynômes P et Q respectivement comme des indéterminées, afin d'obtenir l'expression du résultant en fonction des ces racines. Ceci veut dire précisément que l'on considère les polynômes

$$P = u_0(X-\alpha_1)\cdots(X-\alpha_p) \qquad Q = v_0(X-\beta_1)\cdots(X-\beta_q),$$

éléments de $\mathbb{Z}[u_0, v_0, \underline{\alpha}, \underline{\beta}][X]$. On a alors $\mathrm{res}(P,Q) = \Phi(u_0, v_0, \underline{\alpha}, \underline{\beta})$. Ce Φ est un polynôme à coefficients entiers que l'on va déterminer.

Lemme 4 *Le polynôme $\Phi(u_0, v_0, \underline{\alpha}, \underline{\beta})$ est divisible par $\alpha_i - \beta_j$ pour tout i compris entre 1 et p et tout j compris entre 1 et q.*

Démonstration : Pour fixer les idées, on prend $i = 1$ et $j = 1$. On fait la division euclidienne de Φ par $\alpha_1 - \beta_1$, par rapport à l'indéterminée α_1. On obtient

$$\Phi(u_0, v_0, \underline{\alpha}, \underline{\beta}) = (\alpha_1 - \beta_1)S(u_0, v_0, \underline{\alpha}, \underline{\beta}) + T(u_0, v_0, \alpha_2, \ldots, \alpha_p, \underline{\beta}). \qquad \dagger$$

On spécialise les indéterminées $u_0, v_0, \underline{\alpha}, \underline{\beta}$ en choisissant $a_0, b_0, \underline{\lambda}, \underline{\mu}$ dans $\mathbb{C}^{2+p+q}$ avec $\lambda_1 = \mu_1$. Alors P et Q se spécialisent en des polynômes de $\mathbb{C}[X]$ qui s'annulent tous les deux en $\lambda_1 = \mu_1$, et donc leur résultant $\Phi(a_0, b_0, \underline{\lambda}, \underline{\mu})$ est nul. En reportant dans l'égalité †, on trouve $T(a_0, b_0, \lambda_2, \ldots, \lambda_p, \underline{\mu}) = 0$, et ceci quel que soit le choix de $a_0, b_0, \lambda_2, \ldots, \lambda_p, \underline{\mu}$ dans $\mathbb{C}^{1+p+q}$. Donc T est le polynôme nul et $\alpha_1 - \beta_1$ divise Φ. □

Théorème 5 *Avec les notations ci-dessus, on a*

$$\mathrm{res}(P, Q) = u_0^q v_0^p \prod_{i=1}^{p} \prod_{j=1}^{q} (\alpha_i - \beta_j) = u_0^q \prod_{i=1}^{p} Q(\alpha_i) = (-1)^{pq} v_0^p \prod_{j=1}^{q} P(\beta_j).$$

Démonstration : Les $\alpha_i - \beta_j$ sont premiers entre eux deux à deux et ils divisent tous Φ. Par factorialité de $\mathbb{Z}[u_0, v_0, \underline{\alpha}, \underline{\beta}, X]$, leur produit $\prod_{i,j}(\alpha_i - \beta_j)$ divise Φ. On a $\Phi = F \prod_{i,j}(\alpha_i - \beta_j)$, et on cherche à identifier F. On remarque que

$$Q = v_0 X^q - v_0 \sigma_1(\underline{\beta}) X^{q-1} + \cdots + (-1)^q v_0 \sigma_q(\underline{\beta}),$$

où les σ_j sont les polynômes symétriques élémentaires. En particulier, chaque σ_j est homogène de degré j en $\underline{\beta}$. L'inspection du déterminant de la matrice de Sylvester qui calcule Φ montre que la partie homogène de plus haut degré en les $\underline{\beta}$ de Φ est donnée par la diagonale, et vaut $u_0^q v_0^p (-1)^{pq} (\beta_1 \cdots \beta_q)^p$. Comme la partie homogène de plus haut degré en les $\underline{\beta}$ de $\prod_{i,j}(\alpha_i - \beta_j)$ est $(-1)^{pq}(\beta_1 \cdots \beta_q)^p$, on en déduit que $F = u_0^q v_0^p$, ce qui donne le résultat annoncé. □

Exercice 10 Comparer le résultant de $P(X + a)$ et $Q(X + a)$ avec celui de P et Q.

Exercice 11 Montrer que $\mathrm{res}(P, Q_1 Q_2) = \mathrm{res}(P, Q_1)\mathrm{res}(P, Q_2)$.

La démonstration donnée ci-dessus est celle de [La]. Dans [Ta], on commence en fait par montrer directement le résultat de l'exercice 11 pour établir l'expression du résultant en fonction des racines. Le procédé est intéressant (compléter les détails à titre d'exercice). On part d'un polynôme P de degré p à coefficients dans un corps K et on considère la K-algèbre $A = K[X]/P$. Elle est de dimension p sur K avec une base $\mathcal{B}$ formée des classes de $1, X, \ldots, X^{p-1}$. Soit Q un polynôme de degré q de $K[X]$. La multiplication par la classe de Q est un endomorphisme K-linéaire de A, noté ψ_Q. Les coordonnées dans $\mathcal{B}$ de $\psi_Q(X^i)$, pour $i = 0, \ldots, p-1$, sont les coefficients du reste R_i de la division euclidienne de $X^i Q$ par P. Par ailleurs, le déterminant de la matrice de Sylvester de P et Q est inchangé si on remplace sa $p + q - i$-ème ligne (celle des coefficients de $X^i Q$) par la ligne des coefficients de R_i. On en déduit que $\mathrm{res}(P, Q) = u_0^q \det(\psi_Q)$. Comme $\psi_{Q_1}\psi_{Q_2} = \psi_{Q_1 Q_2}$, il vient $\mathrm{res}(P, Q_1 Q_2) = \mathrm{res}(P, Q_1)\mathrm{res}(P, Q_2)$.

En particulier, pour P unitaire, $\mathrm{res}(P, Q)$ est le déterminant de ψ_Q. C'est le point de vue utilisé dans [Bo] pour traiter le résultant et établir ses propriétés. Le déterminant de l'endomorphisme de multiplication par un élément f de $A = K[X]/P$ est appelé dans [Bo] la norme de f sur K et noté $\mathrm{N}_{A/K}(f)$ (si $K = \mathbb{R}$ et $P = X^2 + 1$, on trouve le carré de la norme habituelle d'un nombre complexe).

Citons sous forme d'exercice une troisième méthode pour établir l'expression du résultant en fonction des racines ([Mi], [Sa]).

Exercice 12 Soient P et Q deux polynômes de degrés p et q respectivement, à coefficients dans un corps algébriquement clos $\overline{K}$. On a

$$P = a_0(X - \alpha_1) \cdots (X - \alpha_p) \qquad Q = b_0(X - \beta_1) \cdots (X - \beta_q)\ ,$$

où a_0, b_0, les α_i et les β_j sont dans $\overline{K}$. On pose

$$\Psi(P, Q) = a_0^q b_0^p \prod_{i=1}^{p} \prod_{j=1}^{q} (\alpha_i - \beta_j)\ .$$

Montrer : (i) $\Psi(Q,P) = (-1)^{pq}\Psi(P,Q)$; (ii) Si le reste de la division euclidienne de Q par P est le polynôme R de degré d, $\Psi(P,Q) = a_0^{q-d}\Psi(P,R)$; (iii) Si b est une constante, $\Psi(P,b) = b^p$.

En comparant comment se calculent Ψ et le résultant en utilisant l'algorithme d'Euclide (voir l'exercice 3), déduire $\mathrm{res}(P,Q) = \Psi(P,Q)$.

Il est à noter que, dans certains ouvrages, le résultant est défini à partir de son expression en fonction des racines.

L'expression du résultant en fonctions des racines $\underline{\alpha}$ et $\underline{\beta}$ montre que le résultant est homogène de degré pq en $(\underline{\alpha}, \underline{\beta})$. Ceci peut se reformuler de la manière suivante. Si $P = u_0X^p + u_1X^{p-1} + \cdots + u_p$, on attribue à chaque coefficient u_i le poids i (c'est le degré de u_i comme polynôme symétrique homogène en les racines). On fait de même pour $Q = v_0X^q + \cdots + v_q$. On attribue naturellement à chaque monôme $\prod_i u_i^{m_i} \prod_j v_j^{n_j}$ le poids $\sum_i im_i + \sum_j jn_j$. Alors, tous les monômes qui apparaissent dans le résultant $\mathrm{res}(P,Q) \in \mathbb{Z}[\underline{u}, \underline{v}]$ sont de poids pq. On dit que $\mathrm{res}(P,Q)$ est *isobare de poids pq en les* $(\underline{u}, \underline{v})$. L'exercice suivant fournit une application de ceci. Cette application est la clé de démonstrations « à l'ancienne » du fameux théorème de Bezout qui dit que deux courbes algébriques de degré p et q sans composante commune se coupent en pq points (comptés avec multiplicité) dans le plan projectif complexe (on peut voir à ce sujet [Ma], chapitre 11, section 8).

Exercice 13 Soient P et Q deux polynômes homogènes en les variables (X,Y,Z), de degrés p et q respectivement. Montrer que $\mathrm{res}_X(P,Q)$ est un polynôme homogène de degré pq en (Y,Z).

Il y a un autre résultat d'homogénéité du résultant (ne pas confondre), moins intéressant, qui ne fait pas intervenir de poids et qui se montre directement à partir de la définition.

Exercice 14 Montrer que le résultant $\mathrm{res}(P,Q)$ est un polynôme homogène de degré q en les coefficients $\underline{u}$ de P, et homogène de degré p en les coefficients $\underline{v}$ de Q.

4 Discriminant

Soit $P = X^p + a_1X^{p-1} + \cdots + a_p$ un polynôme unitaire à coefficients dans un corps K. Soient $\alpha_1, \ldots, \alpha_p$ les p racines de P (comptées avec multiplicité) dans un corps algébriquement clos $\overline{K}$ contenant K. Le *discriminant* de P, noté $\mathrm{dis}(P)$, est

$$\mathrm{dis}(P) = \prod_{1 \le i < j \le p} (\alpha_i - \alpha_j)^2.$$

Autrement dit, le discriminant de P est le produit des carrés des différences entre ses racines.

Exercice 15 Calculer le discriminant de $X^3 + pX + q$.

Exercice 16 Soit P un polynôme unitaire à coefficients réels, de degré p. Montrer que si $\mathrm{dis}(P) > 0$, alors le nombre de racines réelles de P est congru à p modulo 4, et que si $\mathrm{dis}(P) < 0$, alors le nombre de racines réelles de P est congru à $p-2$ modulo 4. Si $P = X^3 + pX + q$, discuter son nombre de racines réelles.

Théorème 6 *Il existe un unique polynôme à coefficient entiers $D_p(u_1, \ldots, u_p)$ tel que pour tout polynôme unitaire $P = X^p + a_1X^{p-1} + \cdots + a_p$ de degré p à coefficients dans un corps K on ait* $\mathrm{dis}(P) = D_p(a_1, \ldots, a_p)$

Les propriétés suivantes sont équivalentes :

1. *Les polynômes P et P' ont un facteur commun non constant.*
2. *Le polynôme P a une racine multiple dans un corps algébriquement clos $\overline{K}$ contenant K.*
3. $\mathrm{dis}(P) = 0$.

Démonstration : On considère $\prod_{1 \le i < j \le p}(T_i - T_j)^2$ qui est un polynôme à coefficients entiers en les indéterminées $T_1, \ldots, T_p$. C'est un polynôme symétrique en les T_i. Donc il s'écrit, et de manière unique, comme polynôme $\Pi(\sigma_1, \ldots, \sigma_p)$ à coefficients entiers en les polynômes symétriques élémentaires des T_i. On pose alors $D_p(u_1, \ldots, u_p) = \Pi(-u_1, u_2, \ldots, (-1)^p u_p)$. On a bien, pour tout corps K et tout polynôme

unitaire $P = X^p + a_1X^{p-1} + \cdots + a_p$ de degré p dans $K[X]$, $\mathrm{dis}(P) = D_p(a_1, \ldots, a_p)$. Noter que le discriminant de P appartient bien à K, et ne dépend pas du choix de l'extension algébriquement close $\overline{K}$.

Montrons maintenant l'équivalence annoncée. L'équivalence des deux premières propriétés a été vue à la proposition 1. L'équivalence des deux dernières est claire avec la définition donnée pour le discriminant. □

Exercice 17 On note $M_p(\mathbb{C})$ l'espace des matrices carrées $p \times p$ à coefficients dans $\mathbb{C}$. Montrer que le sous-ensemble des matrices qui ont toutes leurs valeurs propres distinctes est ouvert dans $M_p(\mathbb{C})$. Est-ce qu'il en est de même pour les matrices diagonalisables ?

On n'est pas, pour définir le discriminant, parti d'un polynôme unitaire $X^p + u_1X^{p-1} + \ldots + u_p$ à coefficients $u_1, \ldots, u_p$ indéterminés. La première partie du théorème 6 permet de le faire, en disant que le discriminant d'un tel polynôme est $D_p(u_1, \ldots, u_p)$, élément de $\mathbb{Z}[u_1, \ldots, u_p]$. On peut ensuite spécialiser $u_1, \ldots, u_p$.

Exercice 18 ([LeAr]) On connaît le déterminant de Vandermonde

$$V(\alpha_1, \ldots, \alpha_p) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_p \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_p^2 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{p-1} & \alpha_2^{p-1} & \ldots & \alpha_p^{p-1} \end{vmatrix}.$$

Montrer que, si $\alpha_1, \ldots, \alpha_p$ sont les racines de P, on a $\mathrm{dis}(P) = V(\alpha_1, \ldots, \alpha_p)^2$. En déduire que

$$\mathrm{dis}(P) = \begin{vmatrix} \nu_0 & \nu_1 & \nu_2 & \ldots & \nu_{p-1} \\ \nu_1 & \nu_2 & \iddots & \iddots & \nu_p \\ \nu_2 & \iddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ \nu_{p-1} & \nu_{p-2} & \ldots & \ldots & \nu_{2p-2} \end{vmatrix},$$

où les ν_k sont les *sommes de Newton* des racines α_i définies par $\nu_0 = p$ et $\nu_k = \sum_{i=1}^p \alpha_i^k$ pour $k > 0$.

Les équivalences du théorème 6 et celles du théorème 3 montrent que le discriminant du polynôme unitaire P est nul si et seulement si le résultant de P et de P' est nul. On a en fait :

Proposition 7 *Soit P un polynôme unitaire de degré p en X. Alors*

$$\mathrm{dis}(P) = (-1)^{p(p-1)/2}\mathrm{res}(P, P').$$

Démonstration : L'expression du résultant en fonction des racines nous donne

$$\begin{aligned} \mathrm{res}(P, P') &= \prod_{i=1}^p P'(\alpha_i) = \prod_{i=1}^p \prod_{j \neq i} (\alpha_i - \alpha_j) = (-1)^{p(p-1)/2} \prod_{1 \leq i < j \leq p} (\alpha_i - \alpha_j)^2 \\ &= (-1)^{p(p-1)/2}\mathrm{dis}(P). \end{aligned}$$

□

Il y a ce signe embêtant qui distingue le discriminant (dans sa définition traditionnelle) et le résultant de P et P'. Ceci peut conduire à une erreur de signe comme dans la première édition de [La]. D'autres ouvrages (comme [Ta]) choisissent de définir le discriminant comme le résultant de P et P'.

Nous n'avons pour le moment défini le discriminant que dans le cas d'un polynôme unitaire. Soit $P = u_0X^p + \cdots + u_p$ un polynôme à coefficients indéterminés, que l'on ne suppose plus unitaire. On se ramène à un polynôme unitaire (à coefficients dans le corps de fractions de $\mathbb{Z}[\underline{u}]$) en divisant par u_0, et le discriminant de ce polynôme unitaire est $\mathrm{dis}(P/u_0) = (-1)^{p(p-1)/2}\mathrm{res}(P/u_0, P'/u_0)$ d'après la

proposition 7. Par ailleurs, on a $\mathrm{res}(P,P') = u_0^{2p-1}\mathrm{res}(P/u_0, P'/u_0)$ et, vu que u_0 est en facteur dans la première colonne de la matrice de Sylvester de P et P', $\mathrm{res}(P,P')$ est divisible par u_0 dans $\mathbb{Z}[\underline{u}]$. Donc $u_0^{2p-2}\mathrm{res}(P/u_0, P'/u_0)$ est un polynôme à coefficients entiers en $u_0, \ldots, u_n$. Ceci conduit à poser

$$\mathrm{dis}(P) = u_0^{2p-2} \prod_{1 \le i < j \le p} (\alpha_j - \alpha_i)^2,$$

et on a :

Proposition 8 *Le discriminant* $\mathrm{dis}(P)$ *est un polynôme à coefficients entiers en les coefficients* $u_0, \ldots, u_p$ *de* P*, et*

$$u_0 \,\mathrm{dis}(P) = (-1)^{p(p-1)/2}\mathrm{res}(P, P').$$

Ici aussi, certains auteurs s'écartent de la tradition. Dans [Ta] par exemple, le discriminant est dans tous les cas le résultant de P et P', mais alors la formule de la proposition 5.9 loc.cit. (conforme à la tradition au signe près) est fausse (lire a_p^{2p-1} au lieu de a_p^{2p-2}). On peut noter que le discriminant de Maple est bien le discriminant traditionnel, tel qu'il a été défini ici.

Exercice 19 Calculer le discriminant de $P = aX^2 + bX + c$. Comparer avec $\mathrm{res}(P, P')$.

Exercice 20 Soient P_1 et P_2 deux polynômes unitaires. Exprimer $\mathrm{dis}(P_1P_2)$ en fonction de $\mathrm{dis}(P_1)$, $\mathrm{dis}(P_2)$ et $\mathrm{res}(P_1, P_2)$

L'exercice suivant offre une justification plus sérieuse que la tradition du choix fait dans la définition du discriminant. Notons $K[X]_p$ l'espace des polynômes à coefficients dans K de degré $\le p$. Soit $\mathrm{dis}_p : K[X]_p \to K$ la fonction polynomiale en les coefficients $u_0, \ldots, u_p$ qui donne le discriminant d'un polynôme (Proposition 8) ; on l'applique même quand un certain nombre des coefficients de tête s'annulent. On considère aussi l'action (à droite) du groupe spécial linéaire $\mathrm{SL}(2,K)$ sur $K[X]_p$ définie par :

$$P(X) \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (cX+d)^p \, P\left(\frac{aX+b}{cX+d}\right) .$$

Exercice 21 Montrer que dis_p est un invariant pour l'action de $\mathrm{SL}(2,K)$: pour tout $P \in K[X]_p$ et tout $g \in \mathrm{SL}(2,K)$, $\mathrm{dis}_p(P \cdot g) = \mathrm{dis}_p(P)$. (On peut choisir des générateurs simples de $\mathrm{SL}(2,K)$, et supposer pour commencer que P est effectivement de degré p et à terme constant non nul.)

5 Une démonstration du théorème des zéros utilisant le résultant

Cette section s'inspire des toutes premières pages du volume 2 de Modern Algebra de van der Waerden. On y voit comment le résultant peut être utilisé pour éliminer une variable dans un système de plus de deux équations. La démonstration du théorème des zéros de Hilbert donnée ci-dessous est terriblement vieux jeu. Elle présente l'avantage d'être effective, en ce sens qu'on peut en extraire un algorithme pour calculer les polynômes $h_1, \ldots, h_m$ de l'énoncé (comparer avec les exercices 5.20 et 5.21 dans le recueil de S. Francinou et H. Gianella).

Lemme 9 *Soit* K *un corps infini,* $P \in K[X_1, \ldots, X_n]$ *un polynôme non nul de degré* d*. Il existe* $(a_1, \ldots, a_{n-1})$ *dans* K^{n-1} *tel que le polynôme*

$$P(X_1 + a_1X_n, \ldots, X_{n-1} + a_{n-1}X_n, X_n)$$

soit de la forme $cX_n^d + Q$*, où* c *est un élément non nul de* K *et* Q *un polynôme de degré* $< d$ *par rapport à* X_n*.*

Démonstration : Soit Π la partie homogène de degré d de P. On a

$$P(X_1 + a_1X_n, \ldots, X_{n-1} + a_{n-1}X_n, X_n) = \Pi(a_1, \ldots, a_{n-1}, 1)X_n^d + Q \ ,$$

où Q est un polynôme de degré $< d$ par rapport à X_n. Puisque le polynôme $\Pi(T_1, \ldots, T_{n-1}, 1)$ n'est pas nul et que K est infini, on peut choisir $(a_1, \ldots, a_{n-1})$ dans K^{n-1} pour que $\Pi(a_1, \ldots, a_{n-1}, 1) \neq 0$. □

Théorème 10 (Théorème des zéros de Hilbert) *Soit K un corps algébriquement clos, et soient $f_1, \dots, f_m$ des polynômes de $K[X_1, \dots, X_n]$ sans zéro commun dans K^n. Alors il existe des polynômes $h_1, \dots, h_m$ de $K[X_1, \dots, X_n]$ tels que $1 = f_1 h_1 + \cdots + f_m h_m$.*

Démonstration : On procède par récurrence sur n. Pour $n = 1$, l'idéal engendré par $f_1, \dots, f_m$ dans $K[X_1]$ est principal, et engendré par g. Si g n'est pas constant, il a un zéro dans K puisque K est algébriquement clos, et ce zéro est commun à tous les f_i.

Supposons maintenant $n > 1$, et le théorème vrai pour $n-1$. On peut supposer qu'aucun des f_i n'est nul. Puisque K est infini, on peut supposer que le polynôme f_1 est unitaire en X_n, quitte à faire un changement linéaire de variables (Lemme 9). Introduisons une nouvelle indéterminée U, et posons

$$g(U, X_1, \dots, X_n) = f_2 + U f_3 + \cdots + U^{m-2} f_m.$$

On calcule le résultant de f_1 et de g par rapport à X_n. Ce résultant appartient à $K[U, X_1, \dots, X_{n-1}]$, et on l'écrit

$$\mathrm{res}_{X_N}(f_1, g) = D_k(X_1, \dots, X_{n-1})U^k + \cdots + D_0(X_1, \dots, X_{n-1}).$$

Ce résultant est dans l'idéal engendré par f_1 et g (proposition 2), et on a des polynômes Λ et Θ de $K[U, X_1, \dots, X_n]$ tels que

$$\mathrm{res}_{X_N}(f_1, g) = \Lambda f_1 + \Theta g.$$

En identifiant les coefficients dans cette égalité entre deux polynômes en U, on voit que les $D_0, \dots, D_k$ sont dans l'idéal engendré par $f_1, \dots, f_m$.

Supposons alors que $D_0, \dots, D_k$ ont un zéro commun x' dans K^{n-1}. Pour tout a appartenant à K, on a $\mathrm{res}_{X_N}(f_1, g)(a, x') = 0$. Comme f_1 est unitaire en X_n, son coefficient dominant en X_n ne s'annule jamais, et donc (théorème 3) l'annulation du résultant entraîne que pour tout $a \in K$ les polynômes $f_1(x', X_n)$ et $g(a, x', X_n)$ ont une racine commune dans K. Comme $f_1(x', X_n)$ a un nombre fini de racines dans K, il y en a une, disons α, qui est racine de $g(a, x', X_n)$ pour une infinité de $a \in K$. Puisque $g(U, x', \alpha)$ a une infinité de racines en U, il est nul, et donc on a $f_2(x', \alpha) = \ldots = f_m(x', \alpha) = 0$. Ainsi (x', α) est un zéro commun à $f_1, \dots, f_m$, ce qui est contraire à l'hypothèse.

On sait donc que $D_0, \dots, D_k$ n'ont aucun zéro commun dans K^{n-1}. Par l'hypothèse de récurrence, 1 appartient à l'idéal engendré par $D_0, \dots, D_k$ dans $K[X_1, \dots, X_{n-1}]$. Comme on a vu que les $D_0, \dots, D_k$ sont dans l'idéal engendré par $f_1, \dots, f_m$ dans $K[X_1, \dots, X_n]$, on en conclut que 1 appartient aussi à cet idéal, ce qui veut dire qu'il existe des polynômes $h_1, \dots, h_m$ de $K[X_1, \dots, X_n]$ tels que $1 = f_1 h_1 + \cdots + f_m h_m$. □

Références

[Bo] N. Bourbaki : *Algèbre. Chapitres 4 à 7*, Masson, 1981.

[La] S. Lang : *Algebra*, Addison-Wesley, 1984.

[LeAr] J. Lelong-Ferrand et J.-M. Arnaudiès : *Cours de Mathématiques. Tome 1 :Algèbre*, Dunod, 1978.

[Ma] M.-P. Malliavin : *Algèbre commutative*, Masson, 1984.

[Mi] M. Mignotte : *Mathématiques pour le calcul formel*, P.U.F., 1981.

[Qu] M. Queysanne : *Algèbre*, Armand Colin, 1964.

[Sa] P. Saux Picart : *Cours de calcul formel. Algorithmes fondamentaux*, Ellipses, 1999.

[Ta] P. Tauvel : *Mathématiques Générales pour l'Agrégation*, Masson, 1992.

Les solutions des exercices calculatoires (traités avec Maple) peuvent être vues à l'adresse

`http://www.maths.univ-rennes1.fr/~coste/exos/exoselim.html`

Toutes les remarques sont les bienvenues pour la prochaine version !