

Le théorème de Bézout et le résultant de deux polynômes

par

Michel Waldschmidt
Université P. et M. Curie (Paris VI)

Introduction.

Soient K un corps algébriquement clos, F_1 et F_2 deux polynômes homogènes de $K[T, X, Y]$ de degrés d_1, d_2 respectivement, sans facteur irréductible commun dans cet anneau factoriel. Nous allons voir que les deux courbes projectives planes $C_1 = Z(F_1)$ et $C_2 = Z(F_2)$ n'ont qu'un nombre fini de points communs, disons P_1, \dots, P_k , et que $k \leq d_1 d_2$. Nous définirons ensuite la multiplicité d'intersection $m(P_i; C_1, C_2)$ de C_1 et C_2 au point P_i , ($1 \leq i \leq k$), et nous montrerons

$$\sum_{i=1}^k m(P_i; C_1, C_2) = d_1 d_2.$$

Enfin nous définissons la multiplicité $m(P, C)$ d'un point P sur une courbe plane C , et nous montrons

$$m(P_i; C_1, C_2) \geq m(P_i, C_1)m(P_i, C_2).$$

Le cas le plus simple est l'intersection d'une courbe affine plane $C_1 \subset \mathbf{A}_2(K)$ d'équation $F(X, Y) = 0$, où $F \in K[X, Y]$ a un degré total d_1 , avec une courbe affine plane C_2 dont l'équation a la forme $Y = Q(X)$, où $Q \in K[X]$ est de degré d_2 . On trouve les coordonnées des points d'intersection en substituant $Q(X)$ à Y dans l'équation de C_1 et en résolvant $F(X, Q(X)) = 0$. Ce cas très simple permet déjà de traiter l'intersection d'une courbe plane quelconque avec une droite ou avec une conique (on peut écrire une conique sous forme $Y = Q(X)$, avec Q de degré 2). Cet exemple montre la nécessité de se placer dans l'espace projectif, et sur un corps algébriquement clos, pour espérer obtenir une égalité dans le théorème de Bézout. Le rôle du résultant est d'*éliminer* la variable Y , même quand l'équation de C_2 n'est pas de la forme $Y = Q(X)$.

§1. Première forme du théorème de Bézout: $k \leq d_1 d_2$.

Nous allons montrer que deux courbes projectives planes C_1, C_2 , de degrés d_1 et d_2 , sans composantes communes, n'ont qu'un nombre fini de points d'intersection, et ce nombre est majoré par le produit $d_1 d_2$. La démonstration utilisera le résultant de deux polynômes.

a) Résultant de deux polynômes en une variable.

Soit A un anneau commutatif unitaire. On désigne par S l'anneau $A[X]$ des polynômes en une variable à coefficients dans A , et, pour d entier ≥ 0 , on note S_d le A -module des polynômes de degré $\leq d$. Ainsi S_d est libre sur A , de rang $d + 1$, une base étant donnée par X^i , ($0 \leq i \leq d$).

Soient P et Q deux polynômes de S de degrés p et q :

$$P(X) = a_0 + a_1 X + \cdots + a_p X^p, \quad Q(X) = b_0 + b_1 X + \cdots + b_q X^q.$$

L'homomorphisme de A -modules

$$\begin{array}{ccc} S_{q-1} \times S_{p-1} & \longrightarrow & S_{p+q-1} \\ (U, V) & \longmapsto & UP + VQ \end{array}$$

a pour matrice, dans les bases citées,

$$\begin{pmatrix} a_0 & 0 & \cdot & \cdot & \cdot & 0 & b_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdot & \cdot & \cdot & 0 & b_1 & b_0 & \cdots & 0 \\ \vdots & \vdots & \cdot & \cdot & \cdot & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{p-1} & a_{p-2} & \cdot & \cdot & \cdot & 0 & b_{p-1} & b_{p-2} & \cdots & b_0 \\ a_p & a_{p-1} & \cdot & \cdot & \cdot & 0 & b_p & b_{p-1} & \cdots & b_1 \\ 0 & a_p & \cdot & \cdot & \cdot & 0 & b_{p+1} & b_p & \cdots & b_2 \\ \vdots & \vdots & \cdot & \cdot & \cdot & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdot & \cdot & \cdot & a_0 & b_{q-1} & b_{q-2} & \cdots & b_{q-p} \\ 0 & 0 & \cdot & \cdot & \cdot & a_1 & b_q & b_{q-1} & \cdots & b_{q-p+1} \\ 0 & 0 & \cdot & \cdot & \cdot & a_2 & 0 & b_q & \cdots & b_{q-p+2} \\ \vdots & \vdots & \cdot & \cdot & \cdot & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdot & \cdot & \cdot & a_p & 0 & 0 & \cdots & b_q \end{pmatrix}$$

Les q premières colonnes sont les composantes, dans la base $(1, X, \dots, X^{p+q-1})$, de $P, XP, \dots, X^{q-1}P$, tandis que les p dernières sont les composantes, dans la même base, de $Q, XQ, \dots, X^{p-1}Q$. La diagonale principale est $(a_0, \dots, a_0, b_q, \dots, b_q)$.
Définition. Le déterminant de cette matrice est appelé le *résultant* de P et Q . On le note $\text{Res}(P, Q)$. Le *résultant universel* est le résultant des deux polynômes

$$U_0 + U_1 X + \cdots + U_p X^p, \quad \text{et} \quad V_0 + V_1 X + \cdots + V_q X^q,$$

dans l'anneau $A_{pq} = \mathbf{Z}[U_0, U_1, \dots, U_p, V_0, V_1, \dots, V_q]$ des polynômes à coefficients dans \mathbf{Z} en $p + q + 2$ indéterminées. On obtient le résultant de P et Q par *spécialisation*, c'est-à-dire comme image par l'homomorphisme canonique de A_{pq} dans A qui envoie U_i sur a_i et V_j sur b_j . Cet homomorphisme canonique n'est injectif que si l'anneau A est de caractéristique nulle.

L'écriture du résultant sous forme de déterminant donne facilement:

Propriété. – *Le résultant universel est un polynôme en $U_0, U_1, \dots, U_p, V_0, V_1, \dots, V_q$, homogène de degré q en U_0, \dots, U_p , et homogène de degré p en V_0, \dots, V_q .*

On obtient aussi facilement:

Propriété. – *Il existe deux polynômes U et V dans S , de degrés $< q$ et $< p$ respectivement, tels que le résultant $R = \text{Res}(P, Q)$ de P et Q s'écrive $R = UP + VQ$.*

On en déduit que si P et Q ont un zéro commun (dans A , ou dans un corps contenant A), alors $\text{Res}(P, Q) = 0$. Nous allons voir la réciproque. Nous aurons besoin du résultat suivant:

Propriété. Soient A_0 un anneau, $A = A_0[Y_1, \dots, Y_n]$ l'anneau des polynômes en n indéterminées à coefficients dans A_0 , et P, Q des polynômes de $A_0[Y_0, \dots, Y_n]$, homogènes de degrés p et q respectivement. On considère P et Q comme des éléments de $A[Y_0]$, et on note $R = \text{Res}_{Y_0}(P, Q) \in A$ leur résultant (par rapport à la variable Y_0). Alors R est homogène de degré pq en Y_1, \dots, Y_n .

Démonstration. Ecrivons

$$P = a_0 + a_1Y_0 + \dots + a_pY_0^p, \quad Q = b_0 + b_1Y_0 + \dots + b_qY_0^q,$$

avec a_i et b_j homogènes de degré $p-i$ et $q-j$ respectivement dans A . Soit $R(Y_1, \dots, Y_n) \in A$ le résultant. On a

$$R(TY_1, \dots, TY_n) = \begin{vmatrix} T^p a_0 & 0 & \dots & 0 & 0 & T^q b_0 & 0 & \dots & 0 & 0 \\ T^{p-1} a_1 & T^p a_0 & \dots & 0 & 0 & T^{q-1} b_1 & T^q b_0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & a_p & T a_{p-1} & 0 & 0 & \dots & b_q & T b_{q-1} \\ 0 & 0 & \dots & 0 & a_p & 0 & 0 & \dots & 0 & b_q \end{vmatrix}$$

On multiplie la première colonne par T^q , la seconde par T^{q-1} , \dots , puis la colonne commençant par $T^q b_0$ par T^p , la suivante par T^{p-1} , \dots . On a ainsi multiplié le déterminant par T^r avec

$$r = \sum_{i=1}^q i + \sum_{j=1}^p j = \frac{q(q+1)}{2} + \frac{p(p+1)}{2}.$$

Dans la i -ème ligne, ($1 \leq i \leq p+q$), on peut mettre en facteur $T^{p+q+1-i}$, et on trouve

$$T^r R(TY_1, \dots, TY_n) = T^s R(Y_1, \dots, Y_n),$$

avec $s = \sum_{i=1}^{p+q} i$, donc

$$s - r = \frac{(p+q)(p+q+1)}{2} - \frac{q(q+1)}{2} - \frac{p(p+1)}{2} = pq,$$

ce qui donne le résultat voulu.

Voici une des propriétés fondamentales du résultant.

Proposition. – Si

$$P(X) = a_0 \prod_{i=1}^p (X - \alpha_i) \quad \text{et} \quad Q(X) = b_0 \prod_{j=1}^q (X - \beta_j),$$

alors

$$\begin{aligned} \text{Res}(P, Q) &= a_0^q b_0^p \prod_{i=1}^p \prod_{j=1}^q (\alpha_i - \beta_j) \\ &= (-1)^{pq} b_0^p \prod_{j=1}^q P(\beta_j) \\ &= a_0^q \prod_{i=1}^p Q(\alpha_i). \end{aligned}$$

Démonstration. Par spécialisation on peut supposer que A est l'anneau des polynômes à coefficients dans \mathbf{Z} en les variables $a_0, b_0, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q$. Dans cet anneau factoriel, $\alpha_i - \beta_j$ est un élément irréductible, qui divise $R = \text{Res}(P, Q)$ (car si on spécialise en $\alpha_i = \beta_j$, alors le résultant est nul). On remarque alors que

$$a_0^q b_0^p \prod_{i=1}^p \prod_{j=1}^q (\alpha_i - \beta_j)$$

est homogène de degré q en les coefficients de P , et de degré p en les coefficients de Q . Il en résulte que cet élément est de la forme cR , avec $c \in \mathbf{Z}$. Le coefficient du monôme $a_0^p b_0^q$ étant 1, on obtient l'égalité annoncée.

Corollaire. Soit K un corps contenant A dans lequel P et Q se décomposent en facteurs de degrés 1. Alors le résultant $\text{Res}(P, Q)$ est nul si et seulement si P et Q ont un zéro commun dans K . En particulier, si l'anneau A est factoriel, alors $\text{Res}(P, Q) = 0$ si et seulement si P et Q ont un facteur irréductible commun.

b) Application aux courbes.

Voici une première forme (faible) du théorème de Bézout. On travaille sur un corps K quelconque.

Théorème 1. – Soient F et G deux formes (polynômes homogènes) de $K[T, X, Y]$, de degrés d_1 et d_2 respectivement, sans facteur irréductible commun. Alors l'ensemble des $(t : x : y) \in \mathbf{P}_2(K)$ tels que $F(t, x, y) = G(t, x, y) = 0$ est fini, avec au plus $d_1 d_2$ éléments.

Démonstration. Le principe de la démonstration est le suivant: on prend k points (distincts) communs aux deux courbes $Z(F)$ et $Z(G)$, disons $P_i = (t_i : x_i : y_i)$, ($1 \leq i \leq k$). On choisit des coordonnées homogènes de telle sorte que $(1 : 0 : 0)$ ne soit pas l'un des P_i , ce qui permet de définir la projection $\pi(P_i) = (t_i : x_i) \in \mathbf{P}_1$ des P_i , et on demande en plus que ces projections soient deux-à-deux distinctes. On prend ensuite le résultant de F et G par rapport à Y ; c'est un polynôme non nul, homogène en T, X de degré $d_1 d_2$, qui s'annule en chaque $\pi(P_i)$. D'où la majoration annoncée: $k \leq d_1 d_2$.

Précisons comment se fait le choix des coordonnées homogènes. On considère les droites joignant les points P_i . Comme on peut agrandir le corps K sans affaiblir l'énoncé, on peut choisir un point P_0 en dehors de la réunion de ces droites. On prend un repère projectif tel que ce point ait pour coordonnées projectives $(0 : 0 : 1)$. Le fait que P_0, P_i, P_j ne soient pas alignés pour $i \neq j$ signifie précisément que les deux points $(t_i : x_i)$ et $(t_j : x_j)$ de \mathbf{P}_1 sont distincts.

Exercices.

- Soient C_1, \dots, C_s des courbes affines planes de degré d , sur un corps K algébriquement clos. On suppose que le sous-ensemble $C_1 \cap \dots \cap C_s$ de $\mathbf{A}_2(K)$ est fini. Montrer que cet ensemble a au plus d^2 éléments.
- Soient F_1, \dots, F_s des polynômes de $K[X, Y]$, dont le degré en X est $\leq L$, et le degré en Y est $\leq M$. On note C_i la courbe affine $Z(F_i)$, ($i = 1, \dots, s$), et on suppose que $C_1 \cap \dots \cap C_s$ est fini. Soient $(x_1, y_1), \dots, (x_k, y_k)$ des points distincts de $C_1 \cap \dots \cap C_s$, avec x_1, \dots, x_k deux-à-deux distincts. Montrer que l'on a $k \leq 2LM$.

§2. Multiplicité d'intersection de deux courbes planes.

Le théorème précédent donne seulement une inégalité: $k \leq d_1 d_2$. La raison est claire: on a utilisé le fait qu'un polynôme en une variable de degré d a au plus d racines dans un corps K . Pour obtenir une égalité, il faut d'une part supposer le corps algébriquement clos, et d'autre part compter les racines avec multiplicités. Ceci va nous fournir une des définitions possibles de la multiplicité d'intersection de deux courbes en un point.

a) *Définition de $m(P; C_1, C_2)$.*

Soient C_1 et C_2 deux courbes sur un corps algébriquement clos, et P un point d'intersection de C_1 et C_2 . On va définir un entier $m(P; C_1, C_2)$, qui est la multiplicité d'intersection de C_1 et C_2 au point P . Comme la définition va être locale, on va travailler dans le plan affine. On va choisir des coordonnées (c'est là toute la difficulté: vérifier que la définition ne dépend pas de ce choix), de telle sorte que $P = (0, 0)$; on écrit les équations des deux courbes $F(X, Y) = 0$ et $G(X, Y) = 0$, on désigne par $R(X)$ le résultant par rapport à Y de F et G , et on considère l'ordre du zéro de R au point 0 . Appelons-le m . Il est facile de voir que, dès que l'intersection comporte plus d'un point, m dépend du choix des coordonnées affines. Par définition, $m(P; C_1, C_2)$ sera le minimum de ces valeurs de m pour tous les choix possibles de coordonnées affines avec $P = (0, 0)$.

Montrons déjà que l'entier m est invariant quand on effectue un changement de coordonnées de la forme

$$X' = X \quad Y' = Y + \lambda X,$$

avec $\lambda \in K$. Pour cela considérons le polynôme

$$R(\lambda, X) = \text{Res}_Y(F(X, \lambda X + Y), G(X, \lambda X + Y)).$$

Montrons qu'il ne dépend pas de λ (*). Pour cela écrivons-le

$$R(\lambda, X) = c_0 + c_1 \lambda + \dots + c_N \lambda^N,$$

avec $c_i \in K[X]$, et $c_N \neq 0$. Par hypothèse les deux polynômes F et G sont sans facteur irréductible commun, donc

$$\text{Res}_Y(F, G) = R(0, X) = c_0(X)$$

n'est pas nul. Soit $\alpha \in K$ tel que $c_0(\alpha)c_N(\alpha) \neq 0$. Si on avait $N > 0$, on pourrait trouver une racine λ_0 au polynôme $R(\lambda, \alpha)$ (le corps K est algébriquement clos). Alors les deux polynômes $F(\alpha, \lambda_0 \alpha + Y)$ et $G(\alpha, \lambda_0 \alpha + Y)$ ont un résultant nul, donc une racine commune, disons $Y = \beta$, ce qui entraîne que $F(\alpha, Y)$ et $G(\alpha, Y)$ ont aussi une racine commune, à savoir $\lambda_0 \alpha + \beta$. Ceci contredit le choix de α avec $R(0, \alpha) = c_0(\alpha) \neq 0$.

Il reste à voir l'effet d'un changement de variables de la forme

$$X' = X + \mu Y \quad Y' = Y.$$

On définit maintenant

$$\bar{R}(\mu, X) = \text{Res}_Y(F(X + \mu Y, Y), G(X + \mu Y, Y)).$$

C'est encore un polynôme en μ et X ; écrivons-le sous la forme:

$$\bar{R}(\mu, X) = A_m(\mu)X^m + \dots + A_N(\mu)X^N,$$

avec $m \leq N$ et $A_m \neq 0$. Alors pour tous les μ pour lesquels $A_m(\mu) \neq 0$, l'ordre de $\bar{R}(\mu, X)$ au point $X = 0$ est égal à m , et pour les autres μ l'ordre en question est plus grand. Donc cet entier m n'est autre que $m(P; C_1, C_2)$.

b) *Le théorème de Bézout (forme définitive)*

(*) Cela résulte aussi de la proposition du §1

Théorème 2. – Soient C_1 et C_2 deux courbes projectives planes, sur un corps algébriquement clos, de degrés d_1 et d_2 respectivement, sans composantes communes. Soient P_1, \dots, P_k leurs points d'intersection. Alors

$$\sum_{i=1}^k m(P_i; C_1, C_2) = d_1 d_2.$$

Démonstration. On reprend la démonstration du théorème 1. On sait que les points d'intersection sont en nombre fini (par le théorème 1). On choisit un système de coordonnées projectives du plan dans lequel ces points ont des coordonnées $(t_i : x_i : y_i)$ avec $t_i \neq 0$, et ont des projections $\pi(P_i) = (t_i : x_i) \in \mathbf{P}_1$, ($1 \leq i \leq k$) deux-à-deux distinctes. Ces points P_i sont donc dans le complémentaire de l'hyperplan $t_i = 0$, que l'on munit de sa structure de plan affine $\mathbf{A}_2(K)$. Soient $F(X, Y) = 0$ et $G(X, Y) = 0$ les équations correspondantes des courbes affines $C_1 \cap \mathbf{A}_2(K)$ et $C_2 \cap \mathbf{A}_2(K)$. Notons m_i la multiplicité du point x_i/t_i comme zéro du résultant $R(X) = \text{Res}_Y(F, G)$. Ce résultant R est un polynôme en X de degré $d_1 d_2$, et ses racines sont $x_1/t_1, \dots, x_k/t_k$, avec les multiplicités m_1, \dots, m_k . Donc

$$m_1 + \dots + m_k = d_1 d_2.$$

Rappelons que m_i dépend du choix des coordonnées, que $m_i \leq m(P_i; C_1, C_2)$, et que l'égalité a lieu pour presque tout choix des coordonnées ; plus précisément, pour tout choix "générique" de coordonnées, (c'est-à-dire sur un ouvert de Zariski), on $m_i = m(P_i; C_1, C_2)$, ce qui donne le résultat annoncé. On obtient de plus $m_i = m(P_i; C_1, C_2)$ pour tout choix de coordonnées dans lequel les x_i sont deux-à-deux distincts.

§3. Multiplicité d'un point sur une hypersurface.

Soit C une hypersurface projective dans $\mathbf{P}_n(K)$, et P un point de C . On va définir la multiplicité $m(P, C)$ de P sur C . On choisit un hyperplan projectif ne contenant pas P , puis un repère affine du complémentaire $\mathbf{A}_n(K)$ de cet hyperplan dans lequel P a pour coordonnées $(0, \dots, 0)$. On écrit l'équation de $C \cap \mathbf{A}_n(K)$ sous la forme $F(X_1, \dots, X_n) = 0$, avec $F \in K[X_1, \dots, X_n]$. On écrit alors F comme somme de polynômes homogènes :

$$F(X_1, \dots, X_n) = F_m(X_1, \dots, X_n) + \dots + F_d(X_1, \dots, X_n),$$

avec $m \leq d$, F_i de degré d_i , ($m \leq i \leq d$) et $F_m \neq 0$. Comme $F(0) = 0$, on a $m \geq 1$. Cet entier m ne dépend pas du choix des coordonnées choisies ; on le note $m(P, C)$. Le point P est dit *simple* (ou *régulier*) sur C si $m = 1$; dans ce cas $F_1(X_1, \dots, X_n) = 0$ est l'équation d'un hyperplan affine, appelé *hyperplan tangent* à C au point P .

Proposition. – Soient C_1 et C_2 deux courbes projectives planes, et soit $P \in C_1 \cap C_2$. Alors

$$m(P; C_1, C_2) \geq m(P, C_1)m(P, C_2).$$

Démonstration. Il s'agit de vérifier que si F et G sont deux éléments de $K[X, Y]$, s'écrivant sous la forme

$$F = F_r + \dots + F_{d_1}, \quad G = G_s + \dots + G_{d_2}$$

avec F_i et G_j homogènes de degrés i et j respectivement, et $r \leq d_1$, $s \leq d_2$, alors leur résultant $R(X) = \text{Res}_Y(F, G)$ a un zéro à l'origine d'ordre au moins rs . On reprend un argument déjà utilisé au §1 : on écrit

$$F = f_0 X^r + f_1 X^{r-1} Y + \dots + f_r Y^r + \dots, \quad G = g_0 X^s + g_1 X^{s-1} Y + \dots + g_s Y^s + \dots$$

et le résultant s'écrit

$$\begin{pmatrix} f_0 X^r & 0 & \dots & 0 & g_0 X^s & 0 & \dots & 0 \\ f_1 X^{r-1} & f_0 X^r & \dots & 0 & g_1 X^{s-1} & g_0 X^s & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_r & f_{r-1} X & \dots & 0 & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

On multiplie la première colonne par X^s , la seconde par X^{s-1} , ..., puis la colonne commençant par $g_0 X^s$ par X^r , la suivante par X^{r-1} , ... On a ainsi multiplié le résultant par une puissance de X , avec l'exposant

$$\sum_{i=1}^r i + \sum_{j=1}^s j = \frac{r(r+1)}{2} + \frac{s(s+1)}{2}.$$

Dans la i -ème ligne, ($1 \leq i \leq r+s$), on peut mettre en facteur $X^{r+s+1-i}$, donc la multiplicité du zéro à l'origine de R est au moins

$$\sum_{i=1}^{r+s} i - \frac{r(r+1)}{2} - \frac{s(s+1)}{2} = rs.$$

Références.

- S. Lang. Algebra; Third Ed., Addison Wesley, 1993.
Voir Chap. IV, §8, p.200–204 pour la définition et les propriétés de base du résultant; voir aussi Chap. IX, §3 et §4.
- P.Samuel. Géométrie projective; PUF, 1986.
Voir Chap. I, §C, p. 26–29 pour la notion de multiplicité d'un point sur une hypersurface.
- R.J. Walker. Algebraic curves; Springer Verlag, 1978.
Voir Chap. I, §9 et §10 pour le résultant, Chap. III, §2 pour la multiplicité d'un point sur une courbe, et Chap. III, §3 pour une forme du théorème de Bézout (tenant compte du produit des multiplicités des points sur chaque courbe). Voir aussi Chap. IV, §5 pour des compléments.
- G. and M. Orzech. Plane algebraic curves; Marcel Dekker, 1981.
Voir Chap. 18, p.174–178, où la multiplicité d'intersection est définie en termes des anneaux locaux des courbes au point considéré.
- R. Hartshorne. Algebraic geometry; Springer Verlag, Graduate Texts **52** 1977.
Pour tout savoir sur le sujet !