Résultant et discriminant

Daniel PERRIN

Avertissement *Le texte ci-dessous est le second provenant de la récupération de mes vieux papiers du temps de Sèvres (l'école normale supérieure de jeunes filles). L'objectif de ces textes est de compléter le cours d'algèbre [DP].*

1 Introduction et notations

La problématique du résultant est la suivante. On a deux polynômes, à une ou plusieurs variables, et il s'agit de savoir s'ils ont des zéros communs. Typiquement, ce genre de situation se rencontre en géométrie algébrique lorsqu'il s'agit de déterminer l'intersection de deux courbes planes, mais aussi en théorie de Galois et dans bien d'autres cadres.

1.1 Notations. On considère un anneau commutatif A et deux polynômes $P, Q \in A[T]$:

$$P(T) = a_m T^m + \cdots + a_1 T + a_0, \quad Q(T) = b_n T^n + \cdots + b_1 T + b_0,$$

avec $m, n \geq 1$. Contrairement à l'usage courant, on ne suppose pas ici a_m et b_n non nuls *a priori*.

2 Rappels sur les déterminants

Pour la commodité du lecteur, nous rappelons ici quelques résultats sur les déterminants que l'on pourra trouver dans tout bon cours d'algèbre linéaire.

2.1 Les définitions

2.1.1 Déterminant

2.1 Proposition-Définition. *Soit M un A-module libre de rang n, muni d'une base $\mathscr{B} = (e_1, \ldots, e_n)$. Il existe une unique forme f, n-linéaire alternée,*

qui vérifie $f(e_1,\dots,e_n)=1$. *Si* $x_1,\dots,x_n$ *est un* n*-uplet de vecteurs de* M, *et si l'on écrit* $x_j=\sum_{i=1}^n x_{ij}e_i$, *on a :*

$$f(x_1,\dots,x_n)=\sum_{\sigma\in\mathfrak{S}_n}\epsilon(\sigma)x_{\sigma(1)1}x_{\sigma(2)2}\cdots x_{\sigma(n)n}=\sum_{\sigma\in\mathfrak{S}_n}\epsilon(\sigma)x_{1\sigma(1)}x_{2\sigma(2)}\cdots x_{n\sigma(n)}.$$

La quantité $f(x_1,\dots,x_n)$ *est appelée* **déterminant** *des* x_j *sur la base* $\mathscr{B}$ *et on note* $f(x_1,\dots,x_n)=\det_{\mathscr{B}}(x_1,\dots,x_n)$.

Démonstration. Il suffit d'appliquer la multilinéarité et l'alternance pour obtenir la première formule et on vérifie inversement que cette formule donne bien une forme n-linéaire alternée. Pour la deuxième version on utilise l'égalité :

$$\epsilon(\sigma)x_{\sigma(1)1}x_{\sigma(2)2}\cdots x_{\sigma(n)n}=\epsilon(\sigma^{-1})x_{1\sigma^{-1}(1)}x_{2\sigma^{-1}(2)}\cdots x_{n\sigma^{-1}(n)}.$$

2.2 Définition. *Soit* $X=(x_{ij})$ *une matrice* $n\times n$ *à coefficients dans* A. *Soit* $M=A^n$ *le module libre canonique de rang* n *et* $\mathscr{B}=(e_1,\dots,e_n)$ *la base canonique de* M *(*$e_1=(1,0,\dots,0)$, *...,* $e_n=(0,\dots,0,1)$*). On pose* $x_j=\sum_{i=1}^n x_{ij}e_i$. *Alors, le déterminant de* X *est, par définition,* $\det_{\mathscr{B}}(x_1,\dots,x_n)$.

2.1.2 Mineurs

2.3 Définition. *Soit* A *un anneau commutatif et soit* X *une matrice* $n\times r$ *à coefficients dans* A. *Pour* $k\leq \mathrm{Min}\,(n,r)$, *on considère des entiers* $1\leq i_1<i_2<\cdots<i_k\leq n$ *et* $1\leq j_1<j_2<\cdots<j_k\leq r$. *On note* $X_{i_1,\dots,i_k;j_1,\dots,j_k}$ *la matrice obtenue en ne conservant dans* X *que les lignes d'indices* i_p *et les colonnes d'indices* j_q. *On pose alors* $\Delta_{i_1,\dots,i_k;j_1,\dots,j_k}=\det X_{i_1,\dots,i_k;j_1,\dots,j_k}$ *et ce déterminant est le* k**-mineur** *de* X *relatif aux lignes d'indices* i_p *et aux colonnes d'indice* j_q.

2.2 Le développement selon une rangée

2.2.1 Les formules

2.4 Proposition. *Soit* M *un* A*-module libre de rang* n, *muni d'une base* $\mathscr{B}=(e_1,\dots,e_n)$, *et soit* $x_1,\dots,x_n$ *un* n*-uplet de vecteurs de* M. *On pose* $x_j=\sum_{i=1}^n x_{ij}e_i$ *et* $d=\det_{\mathscr{B}}(x_1,\dots,x_n)$. *On note* X *la matrice des* x_{ij} *et* X_{ij} *la matrice obtenue*[1] *en supprimant dans* X *la ligne* i *et la colonne* j. *On a les formules (avec les colonnes) :*

1) Pour tout $j=1,\dots,n$, $d=\sum_{i=1}^n x_{ij}(-1)^{i+j}\det(X_{ij})$.

1. Son déterminant est le mineur de X relatif aux lignes d'indices $\neq i$ et aux colonnes d'indices $\neq j$.

2) Pour tous $j, k = 1, \ldots, n$, $j \neq k$, $0 = \sum_{i=1}^{n} x_{ij}(-1)^{i+k} \det(X_{ik})$.

On a aussi les formules analogues avec les lignes :

3) Pour tout $i = 1, \ldots, n$, $d = \sum_{j=1}^{n} x_{ij}(-1)^{i+j} \det(X_{ij})$.

4) Pour tous $i, k = 1, \ldots, n$, $i \neq k$, $0 = \sum_{j=1}^{n} x_{ij}(-1)^{j+k} \det(X_{kj})$.

On peut encore écrire les deux dernières formules en utilisant le symbole de Kronecker δ_{ik} :

$$(*) \qquad \text{pour tous } i, k = 1, \ldots, n, \quad \delta_{ik} d = \sum_{j=1}^{n} x_{ij}(-1)^{j+k} \det(X_{kj}).$$

Démonstration. On applique la définition du déterminant en notant que les termes qui sont en facteur de x_{ij} sont ceux de la matrice X_{ij}. On vérifie ensuite que les signes sont bien ceux annoncés.

2.2.2 Application

2.5 Corollaire. *Soit M un A-module libre de rang n, muni d'une base $\mathscr{B} = (e_1, \ldots, e_n)$, soit $x_1, \ldots, x_n$ un n-uplet de vecteurs de M et soit N le sous-module engendré par les x_j. Posons $d = \det_{\mathscr{B}}(x_1, \ldots, x_n)$. Alors, on a $dM \subset N$. Autrement dit, pour tout i, le vecteur de_i est combinaison linéaire des x_j.*

Démonstration. Si l'on pose $x_j = \sum_{k=1}^{n} x_{ij} e_i$, la formule $(*)$ de la proposition précédente donne, pour tout k : $de_k = \sum_{j=1}^{n} (-1)^{k+j} \det(X_{kj}) x_j$.

2.3 Dépendance

La nullité du déterminant est liée à la dépendance linéaire des vecteurs :

2.6 Théorème. *Soit M un A-module libre de rang n, muni d'une base $\mathscr{B} = (e_1, \ldots, e_n)$, et soit $x_1, \ldots, x_n$ un n-uplet de vecteurs de M. Posons $d = \det_{\mathscr{B}}(x_1, \ldots, x_n)$.*

1) Si d est nul, les x_j sont liés, i.e. il existe $\lambda_1, \ldots, \lambda_n \in A$ non tous nuls tels que $\lambda_1 x_1 + \cdots + \lambda_n x_n = 0$.

2) Inversement, si l'anneau A est intègre et si les x_j sont liés, on a $d = 0$.

Démonstration. Le point 1) est le cas $r = n$ du lemme suivant :

2.7 Lemme. *Soit A un anneau commutatif quelconque, M un A-module libre de rang n muni d'une base $e_1, \ldots, e_n$. Soient $x_1, \ldots, x_r \in M$, avec $r \leq n$, et soit X la matrice $n \times r$ des x_j sur les e_i. On suppose que tous les r-mineurs de X sont nuls. Alors, les x_j sont liés.*

Démonstration. On raisonne par récurrence sur r, le cas $r = 1$ étant évident. Supposons la propriété vraie au rang $r-1$ et passons à r. Si tous les $r-1$-mineurs de X sont nuls, c'est vrai en particulier pour ceux des $r-1$ premières colonnes et, vu l'hypothèse de récurrence, les vecteurs $x_1, \ldots, x_{r-1}$ sont alors liés, donc aussi $x_1, \ldots, x_r$. Sinon, il existe un $r-1$-mineur non nul et on peut supposer qu'il s'agit du mineur porté par les $r-1$ premières lignes et les $r-1$ premières colonnes de X. Notons A_r, A_{r-1}, ..., A_1, les $r-1$-mineurs correspondant aux $r-1$ premières lignes (A_j est le mineur privé de la colonne j). On a donc $A_r \neq 0$. Soient $x_{i1}, \ldots, x_{ir}$ les termes de la i-ième ligne de X. En vertu de 2.4, on a, pour tout $i = 1, \ldots, n$, l'égalité $\sum_{k=1}^{r} x_{ik}(-1)^k A_k = 0$. En effet, pour $i \geq r$, cette égalité exprime que le r-mineur formé des lignes $1, \ldots, r-1, i$ est nul (c'est l'hypothèse) et pour $i \leq r-1$ c'est 2.4.4. Mais alors, on a la relation de dépendance $\sum_{k=1}^{r}(-1)^k A_k x_k = 0$, non triviale puisque A_r est non nul.

2) Supposons que l'on a $\lambda_1 x_1 + \cdots + \lambda_n x_n = 0$, avec, par exemple, $\lambda_1 \neq 0$. On peut plonger A dans son corps des fractions K, ce qui ne change pas le déterminant des x_j. Mais, sur K, λ_1 est inversible et on a $x_1 = -\lambda_1^{-1}\lambda_2 x_2 - \cdots - \lambda_1^{-1}\lambda_n x_n$. En vertu de la multilinéarité et de l'alternance, il en résulte que le déterminant est nul.

2.8 *Remarques.* 1) Si l'anneau est intègre il y a une preuve plus simple du premier point. Quitte à passer au corps des fractions, on peut supposer que A est un corps. Montrons la contraposée de 1) : si la famille $x_1, \ldots, x_n$ est libre, alors on a $d \neq 0$. En effet, $x_1, \ldots, x_n$ est alors une base de M et on peut écrire les e_i sur les x_j avec une matrice $Y = (y_{ij})$. Mais, si f est la forme n-linéaire alternée qui vaut 1 sur les e_i, on a $1 = f(e_1, \ldots, e_n) = (\det Y) f(x_1, \ldots, x_n) = (\det Y)d$ et donc $d \neq 0$.

2) Sur un anneau non intègre, les vecteurs peuvent être liés sans que le déterminant soit nul. Exemple : $A = \mathbf{Z}/6\mathbf{Z}$, $M = A^2$ muni de la base canonique e_1, e_2 ; $x_1 = e_1 + e_2$, $x_2 = e_1 + 3e_2$. Le déterminant vaut $2 \neq 0$, mais on a $3x_1 - 3x_2 = 0$.

3 Définition du résultant et premières propriétés

3.1 La définition

3.1 Définition. *Avec les notations de 1.1, le* **résultant** *de P et Q est l'élément $R(P,Q) \in A$ donné par le déterminant suivant :*

$$R(P,Q) = \begin{vmatrix} a_m & a_{m-1} & \dots & \dots & a_0 & 0 & \dots & 0 & 0 \\ 0 & a_m & a_{m-1} & \dots & \dots & a_0 & \dots & \dots & 0 \\ & & \dots & & & & \dots & & \\ 0 & \dots & 0 & 0 & a_m & a_{m-1} & \dots & \dots & a_0 \\ b_n & b_{n-1} & \dots & \dots & \dots & b_0 & 0 & \dots & 0 \\ 0 & b_n & b_{n-1} & \dots & \dots & \dots & b_0 & \dots & 0 \\ & & \dots & & & & \dots & & \\ 0 & \dots & \dots & 0 & 0 & b_n & b_{n-1} & \dots & b_0 \end{vmatrix}$$

qui est un déterminant $(m+n)\times(m+n)$ avec n lignes formées des coefficients a_i et de $n-1$ zéros et m lignes formées des b_j et de $m-1$ zéros. Précisément, si l'on convient de poser $a_i = 0$ si $i > m$ ou $i < 0$ et de même pour b_j, et si l'on note r_{ij}, $i,j = 1,\dots,m+n$ le terme générique de ce déterminant, on a les formules :

1) pour $1 \le i \le n$, $1 \le j \le m+n$, $r_{ij} = a_{m+i-j}$,

2) pour $n+1 \le i \le m+n$, $1 \le j \le m+n$, $r_{ij} = b_{i-j}$.

3.2 *Remarque.* On notera que le terme diagonal de ce déterminant est égal à $a_m^n b_0^m$.

Pour comprendre le sens de ce déterminant, on considère le A-module M suivant :

$$M = \{F(T) \in A[T] \ \mid \ d^\circ F \le m+n-1.\ \}$$

C'est un A-module libre dont une base est $\mathscr{B} = (T^{m+n-1}, \dots, T, 1)$. On a alors le résultat suivant :

3.3 Proposition. *Le résultant $R := R(P,Q)$ est le déterminant sur la base $\mathscr{B}$ de la famille de polynômes :*

$$T^{n-1}P, T^{n-2}P, \dots, TP, P\,;\, T^{m-1}Q, \dots, TQ, Q.$$

Démonstration. C'est évident en écrivant ces polynômes. Par exemple, on a

$$T^{n-1}P = a_m T^{m+n-1} + a_{n-1}T^{m+n-2} + \dots + a_1 T^n + a_0 T^{n-1},$$

ce qui donne la première ligne de R.

3.2 Le théorème fondamental

3.4 Proposition. *On reprend les notations de 1.1 et 3.1. Il existe $F, G \in A[T]$, non tous deux nuls, tels que l'on ait $R := R(P,Q) = FP + GQ$ avec $d^\circ(F) < n$ et $d^\circ(G) < m$.*

Démonstration. Comme ci-dessus, on considère le sous-A-module M de $A[T]$ formé des polynômes de degré $\leq m+n-1$ et muni de la base canonique. En vertu de 3.3, R est le déterminant de la famille des $T^{n-1}P, T^{n-2}P, \ldots, TP, P$; $T^{m-1}Q, \ldots, TQ, Q$ et, en appliquant 2.5 au polynôme constant 1, on trouve que R est combinaison linéaire des T^iP et T^jQ, donc de la forme $FP + GQ$ avec $d^\circ(F) < n$ et $d^\circ(G) < m$. Cela donne la conclusion si R est non nul (car alors F et G ne sont pas tous deux nuls). Si $R = 0$, c'est le théorème 2.6.

Le cas $R = 0$ peut être précisé comme suit :

3.5 Théorème. *On reprend les notations de 1.1 et 3.1. On considère les deux conditions suivantes :*

1) Le résultant $R := R(P,Q)$ est nul.

2) Il existe $F, G \in A[T]$, non tous deux nuls, tels que $FP + GQ = 0$ avec $d^\circ(F) < n$ et $d^\circ(G) < m$.

Alors, on a 1) $\Longrightarrow$ 2) et la réciproque est vraie si A est intègre. Dans ce cas, si K est le corps des fractions de A et si a_m et b_n sont non nuls, ces conditions sont encore équivalentes aux suivantes :

3) Le pgcd de P et Q dans $K[T]$ est de degré positif.

4) Les polynômes P et Q ont une racine commune dans une extension convenable de K, par exemple $L = D_K(PQ)$.

Démonstration. Le fait que 1) $\Longrightarrow$ 2) résulte de 3.4 et pour la réciproque, on applique 2.6.2 à la famille liée $T^{n-1}P, T^{n-2}P, \ldots, TP, P$; $T^{m-1}Q, \ldots, TQ, Q$.

Montrons que 2) implique 3). Sinon, les polynômes P et Q sont premiers entre eux dans l'anneau factoriel $K[T]$. Comme on a $FP = -GQ$, le théorème de Gauss montre que P divise G, ce qui est absurde pour une raison de degré.

Montrons que 3) implique 2). Soit D le *pgcd* de P, Q dans $K[T]$. On a $P = DG_0$ et $Q = -DF_0$, avec $F_0, G_0 \in K[T]$, de degré plus petits que n et m respectivement puisque D est de degré > 0, donc $F_0P = DF_0G_0 = -G_0Q$. Si e est le produit des dénominateurs des coefficients de F_0 et G_0, les polynômes $F = eF_0$ et $G = eG_0$ sont à coefficients dans A et on a $FP = -GQ$ d'où le point 2).

L'équivalence de 3) et 4) est évidente.

3.6 *Remarque.* Attention, 2) $\Longrightarrow$ 1) n'est pas vrai si A n'est pas intègre. On peut adapter l'exemple de 2.8 : sur $\mathbf{Z}/6\mathbf{Z}$, les polynômes $P(T) = T + 1$ et $Q(T) = T + 3$ ont pour résultant $R = 2 \neq 0$ et pourtant on a $3P - 3Q = 0$.

3.3 Exemples

3.3.1 Le cas $m = n = 1$

On a $P(T) = aT + b$, $Q(T) = cT + d$ et le résultant est égal à $ad - bc$. Si a et c sont non nuls, on a $R = 0$ si les racines $-b/a$ et $-d/c$ sont égales.

3.3.2 Le cas $n = 1$

On suppose $P(T) = a_m T^m + \cdots + a_1 T + a_0$, mais $Q(T) = b_1 T + b_0$. On a donc :

$$R = \begin{vmatrix} a_m & a_{m-1} & \dots & \dots & a_0 & 0 & \dots & 0 \\ b_1 & b_0 & 0 & \dots & \dots & \dots & 0 & \\ 0 & b_1 & b_0 & \dots & & \dots & \dots & 0 \\ & & \dots & & & & & \dots \\ 0 & \dots & \dots & 0 & 0 & 0 & b_1 & b_0 \end{vmatrix}$$

Le développement du déterminant par rapport à sa première ligne donne :

$$R = a_m b_0^m - a_{m-1} b_0^{m-1} b_1 + \cdots + (-1)^i a_{m-i} b_0^{m-i} b_1^i + \cdots + (-1)^m a_0 b_1^m.$$

Dans le cas intègre, avec $b_1 \neq 0$, on voit que R est nul si et seulement si $-b_0/b_1$ est racine de P.

3.3.3 Le cas $m = n = 2$

Un calcul facile donne :

$$R = a_0^2 b_2^2 - a_0 a_1 b_1 b_2 - 2 a_0 a_2 b_0 b_2 + a_0 a_2 b_1^2 + a_1^2 b_0 b_2 - a_1 a_2 b_0 b_1 + a_2^2 b_0^2.$$

4 Résultant et racines

L'objectif de ce paragraphe est de prouver le théorème 4.1 qui explicite la relation entre le résultant et les racines des polynômes.

4.1 Anneaux universels

Lorsqu'on a un anneau A et des éléments $t_1, \dots t_n$ de A, on peut introduire **l'anneau universel** relatif à cette situation qui est l'anneau de polynômes $\Lambda := \mathbf{Z}[T_1, \dots, T_n]$. L'intérêt de cette procédure est que cet anneau universel est factoriel et qu'on va pouvoir y faire des raisonnements de divisibilité. On reviendra à l'anneau A en "spécialisant les T_i" ou encore en "faisant $T_i = t_i$", c'est-à-dire au moyen de l'homomorphisme $\Phi : \Lambda \to A$ qui est défini de

manière canonique sur $\mathbf{Z}$ en associant $n.1_A$ à n et qui à T_i associe t_i. Le fait que Φ soit un homomorphisme signifie que l'on a $\Phi(P(T_1,\dots,T_n)) = P(t_1,\dots,t_n)$ pour tout polynôme P. Cela montre que si l'on a une formule polynomiale dans Λ on en déduit la formule analogue dans A.

En particulier, en ce qui concerne le résultant, on peut remplacer l'anneau de base A par l'anneau "universel" : $\Lambda := \mathbf{Z}[U_m,\dots,U_0,V_n,\dots,V_0]$ où les U_i, V_j sont des indéterminées et considérer les polynômes universels associés : $P(T) = U_m T^m + \cdots + U_1 T + U_0$ et $Q(T) = V_n T^n + \cdots + V_1 T + V_0$ et calculer leur résultant R qui est un polynôme en les U_i, V_j. Les résultats précédents s'appliquent à ce cas et si l'on a un anneau A quelconque et des polynômes P, Q, leur résultant s'obtient en spécialisant le résultant universel, c'est-à-dire en donnant aux indéterminées U_i, V_j des valeurs dans A.

4.2 La formule universelle donnant le résultant en fonction des racines

On va appliquer la procédure de spécialisation dans un autre cas. On considère l'anneau universel "des racines" :

$$\Gamma = \mathbf{Z}[X_1,\dots,X_m,Y_1,\dots,Y_n][U_m,V_n]$$

où les X_i, Y_j, U_m, V_n sont des indéterminées. On a alors le théorème suivant :

4.1 Théorème. *On note $P(T) = U_m(T-X_1)\cdots(T-X_m)$ et $Q(T) = V_n(T-Y_1)\cdots(T-Y_n)$ les polynômes de $\Gamma(T)$. Ce sont les polynômes qui admettent les racines "universelles" X_i et Y_j. Dans l'anneau Γ, on a les formules :*

$$R(P,Q) = U_m^n V_n^m \prod_{i=1}^{m}\prod_{j=1}^{n}(X_i - Y_j),$$

$$R(P,Q) = U_m^n \prod_{i=1}^{m} Q(X_i) = (-1)^{mn} V_n^m \prod_{j=1}^{n} P(Y_j).$$

Démonstration. On commence par un calcul de degrés :

4.2 Lemme. *Avec les notations précédentes, $R := R(P,Q)$ est un polynôme en les variables U_m, V_n, X_i, Y_j qui est le produit de $U_m^n V_n^m$ par un polynôme homogène en les X_i, Y_j de degré total $m+n$.*

Démonstration. Notons a_i (resp. b_j) le coefficient du terme de degré i (resp. j) de P (resp. Q). On a $a_m = U_m$ et, pour $i < m$, $a_i = (-1)^{m-i} U_m \Sigma_{m-i}(X_1,\dots,X_m)$ où Σ_k est le k-ième polynôme symétrique élémentaire en les X_i ($\Sigma_1 = X_1 +$

$\cdots + X_m$, ..., $\Sigma_m = X_1 \cdots X_m$). On sait que Σ_k est homogène de degré k en les X_i, de sorte que a_i est homogène de degré $m-i$ pour $i < m$. De même, on a $b_n = V_n$ et, pour $j < n$, $b_j = (-1)^{n-j} V_n \Sigma_{n-j}(Y_1, \ldots Y_n)$. Le coefficient b_j est homogène de degré $n-j$ en les Y_j pour $j < n$.

Notons r_{ij} le terme générique de R. En vertu de 3.1 on a $r_{ij} = a_{m+i-j}$ pour $i \leq n$ et $r_{ij} = b_{i-j}$ pour $i > n$, avec la convention de nullité explicitée en 3.1. Le déterminant R est donné par la formule :

$$R = \sum_{\sigma \in \mathfrak{S}_{m+n}} \epsilon(\sigma) r_{1,\sigma(1)} \cdots r_{i,\sigma(i)} \cdots r_{m+n,\sigma(m+n)}.$$

L'assertion sur les variables U_m, V_n est claire car les termes des n premières lignes (resp. des m dernières) sont tous le produit de U_m (resp. V_n) par un polynôme en X_i, Y_j. Calculons le degré d'un terme quelconque (non nul) de la somme en les variables X_i, Y_j. Pour $i \leq n$ on a $d^\circ r_{i,\sigma(i)} = d^\circ a_{m+i-\sigma(i)} = \sigma(i) - i$. Pour $i > n$, on a $d^\circ r_{i,\sigma(i)} = d^\circ b_{i-\sigma(i)} = n + \sigma(i) - i$. Comme il y a m termes correspondant à $i > n$, le degré du produit est égal à $mn + \sum_\sigma \sigma(i) - i$, mais, comme σ est une permutation, la somme des $\sigma(i) - i$ est nulle, de sorte que tous les termes sont bien de degré mn.

On continue par un lemme sur les polynômes à coefficients entiers :

4.3 Lemme. *Soit $R \in \mathbf{Z}[T_1, \ldots, T_n]$ un polynôme.*

On suppose que $R(\underline{T}) := R(T_1, \ldots, T_i, \ldots, T_j, \ldots, T_n)$ est nul quand on fait $T_i = T_j$. Alors, $T_i - T_j$ divise R.

Démonstration. On peut effectuer la division euclidienne de R par $T_i - T_j$ relativement à l'indéterminée T_i (voir [DP] II 3.31). On obtient :

$$R(\underline{T}) = (T_i - T_j) S(\underline{T}) + S'(T_1, \ldots \widehat{T_i} \ldots, T_n)$$

où le symbole $\widehat{T_i}$ signifie que l'indéterminée T_i est omise. On sait que R s'annule si l'on fait $T_i = T_j$, ainsi que le polynôme $T_i - T_j$, donc aussi S' qui est donc nul puisque T_i n'intervient pas.

Revenons au théorème.

Soit Γ' l'anneau obtenu à partir de Γ en enlevant la variable Y_j. On considère l'homomorphisme $\Phi : \Gamma \to \Gamma'$ qui à Y_j associe X_i et on note $\overline{P}$ et $\overline{Q}$ les images de P, Q dans Γ'. Comme $\overline{P}$ et $\overline{Q}$ ont une racine commune (la racine $X_i = Y_j$) dans Γ', leur résultant $\overline{R} = \Phi(R)$ est nul. En vertu du lemme, le polynôme $X_i - Y_j$ divise donc R. Dans l'anneau factoriel Γ, les $X_i - Y_j$ sont des irréductibles (car ils sont de degré 1) et ils sont distincts. Comme ils divisent R, leur produit $\prod_{i,j}(X_i - Y_j)$ divise R. On a vu aussi en 4.2 que U_m^n et V_n^m divisent R. En définitive, le polynôme $S := U_m^n V_n^m \prod_{i=1}^m \prod_{j=1}^n (X_i - Y_j)$

divise R. Comme R et S sont homogènes de mêmes degrés en les U_m, V_n, X_i, Y_j, on a $R = \lambda S$ où λ est un entier. Pour calculer λ on regarde le terme diagonal du déterminant qui vaut $U_m^n b_0^m$, soit $U_m^n V_n^m (-1)^{mn} \prod_{j=1}^n Y_j^m$, terme constant de R par rapport aux variables X_i. En faisant $X_i = 0$ dans S, on constate que c'est bien aussi le terme constant de S par rapport aux X_i. On a donc $R = S$ comme annoncé et les autres formules sont immédiates.

4.4 Corollaire. *On reprend les notations de 1.1. Soit A un anneau intègre, K son corps des fractions, $P, Q \in A[T]$ et soit L un corps de décomposition de PQ. On note $x_1, \ldots, x_m$ (resp. $y_1, \ldots, y_n$) les racines de P (resp. Q) dans L, comptées avec d'éventuelles multiplicités. Soit $R = R(P, Q)$. On a les formules :*

$$R(P,Q) = a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x_i - y_j).$$

$$R(P,Q) = a_m^n \prod_{i=1}^{m} Q(x_i) = (-1)^{mn} b_n^m \prod_{j=1}^{n} P(y_j).$$

Démonstration. C'est une conséquence de 4.1 par spécialisation.

5 Applications

5.1 Le théorème de Bézout faible

5.1.1 Bézout

On utilise ici les notations de [IGADP] concernant les courbes algébriques planes. Le résultat en vue est le suivant :

5.1 Théorème. *Soit k un corps, $P, Q \in k[X, Y]$ deux polynômes premiers entre eux, de degrés respectifs m et n. Alors, on a $|V(P) \cap V(Q)| \leq mn$: les courbes affines planes d'équations $P = 0$ et $Q = 0$ ont au plus mn points d'intersection.*

5.2 *Remarque.* Bien entendu, il y a mieux, on peut affirmer qu'on a exactement $|V(P) \cap V(Q)| = mn$, mais il faut pour cela travailler sur un corps algébriquement clos, en projectif, et en comptant des multiplicités, voir [IGADP], VI 2.2.

Le ressort de la preuve est la proposition suivante :

5.3 Proposition. *Soient $P, Q \in k[X, Y]$ des polynômes de degrés (totaux) respectifs m et n et soit $R \in k[X]$ le résultant de ces polynômes vus comme polynômes en Y à coefficients dans $k[X]$. Alors, on a $d^\circ R \leq mn$.*

Démonstration. On écrit $P(Y) = a_k(X)Y^k + \cdots + a_0(X)$ et $Q(Y) = b_l(X)Y^l + \cdots + b_0(X)$ avec $0 \leq k \leq m$ et $0 \leq l \leq n$. Les a_i et b_j sont des polynômes en X dont les degrés sont respectivement $\leq m - i$ et $n - j$ en vertu du fait que les degrés totaux de P et Q sont m et n.

On note (α_{ij}) la matrice dont le déterminant est le résultant et on calcule $R(X)$ à l'aide de la définition du déterminant :

$$R(X) = \sum_{\sigma \in \mathfrak{S}_{k+l}} \epsilon(\sigma) \alpha_{1\sigma(1)} \cdots \alpha_{k+l,\sigma(k+l)}.$$

Calculons les coefficients $\alpha_{i,\sigma(i)}$, d'abord pour $i \leq l$. On a $\alpha_{i,\sigma(i)} = 0$ si $\sigma(i) < i$ et $\alpha_{i,\sigma(i)} = a_{k-\sigma(i)+i}$ si $\sigma(i) \geq i$. En particulier, ces polynômes ont même degré.

Calculons maintenant $\alpha_{i,\sigma(i)}$ pour $i \geq l + 1$. Pour $\sigma(i) > i$ ce coefficient est nul, et sinon, il vaut $b_{i-\sigma(i)}$.

On en déduit qu'on a $d^\circ \alpha_{i,\sigma(i)} \leq m - k + \sigma(i) - i$ pour $i \leq l$ et $d^\circ \alpha_{i,\sigma(i)} \leq n - i + \sigma(i)$ pour $i > l$. Cela permet de majorer le degré de chaque terme :

$$d^\circ \big(\prod_{i=1}^{k+l} \alpha_{i,\sigma(i)} \big) \leq \sum_{i=1}^{l} m - k + \sigma(i) - i + \sum_{i>l} n - i + \sigma(i).$$

Mais, comme σ est une permutation, on a $\sum_{i=1}^{k+l} \sigma(i) - i = 0$ et il reste $d^\circ \big(\prod_{i=1}^{k+l} \alpha_{i,\sigma(i)} \big) \leq l(m-k) + km \leq n(m-k) + kn = mn$.

5.4 Corollaire. *On suppose P et Q premiers entre eux. Alors $Z := V(P) \cap V(Q)$ est fini.*

Démonstration. On considère le résultant $R(X)$ de P et Q vus comme polynômes en Y. On a $R(X) = F(X,Y)P(X,Y) + G(X,Y)Q(X,Y)$, de sorte que, si $(x,y) \in Z$, on a $R(x) = 0$. Il y a donc au plus mn abscisses possibles pour les points d'intersection. De la même manière, on montre qu'il y a au plus mn ordonnées possibles, donc en tout au plus m^2n^2 points.

5.5 *Remarque.* Évidemment, cette majoration de $|Z|$ par m^2n^2 est très mauvaise comme on va le voir.

On peut maintenant prouver 5.1.

Notons d'abord que, quitte à plonger k dans un corps infini (par exemple $k(T)$), on peut supposer k infini.

Comme k est infini et $Z = V(F) \cap V(G)$ fini, il existe deux droites D_1, D_2 distinctes, passant par l'origine du plan k^2 et telles que la droite qui joint deux points quelconques de Z ne soit parallèle ni à D_1 ni à D_2. Si l'on choisit les D_i

comme nouveaux axes, les polynômes obtenus par changement de variables sont encore de degrés m et n. Cette fois, l'ensemble $V(P) \cap V(Q)$ a la propriété qu'il y a au plus un point d'abscisse donnée dans cet ensemble. Comme il y a au plus mn abscisses possibles, on en déduit $|V(P) \cap V(Q)| \leq mn$.

5.1.2 Pascal

5.6 Corollaire. *Soit k un corps, $F, G \in k[X, Y]$ deux polynômes de degré n. On suppose que $V(F)$ et $V(G)$ sont infinis et que l'intersection $Z := V(F) \cap V(G)$ est de cardinal n^2. Soit H un polynôme irréductible de degré $d < n$. On suppose que $V(H)$ est infini et qu'il contient exactement nd points de Z. Alors, les $n(n-d)$ restants sont sur une courbe $V(L)$ avec $d^\circ l = n - d$.*

Démonstration. Soit m un point de $V(H)$ non situé sur Z (un tel point existe car $V(H)$ est infini). Il existe $\lambda, \mu \in k$ tels que $(\lambda F + \mu G)(m) = 0$ (si par exemple $F(m)$ est non nul, on prend $\mu = 1$ et $\lambda = -G(m)/F(m)$). Alors, l'intersection $V(H) \cap V(\lambda F + \mu G)$ contient les nd points de $V(F) \cap V(G) \cap V(H)$, plus le point m. Comme H est de degré d et $\lambda F + \mu G$ de degré n, en vertu de Bézout, cela n'est possible que si H et $\lambda F + \mu G$ ont un facteur commun non trivial. Comme H est irréductible, il divise donc $\lambda F + \mu G$ et on a $\lambda F + \mu G = HL$, avec L de degré $n - d$. Mais alors, les points de Z qui ne sont pas dans $V(H)$ annulent F, G donc $\lambda F + \mu G$, donc L, cqfd.

5.7 Corollaire. (Pascal) *Soit Γ une conique propre non vide de k^2 et soient p, q, r, p', q', r' des points distincts de Γ. Les droites (qr') et $(q'r)$ (resp. (rp') et $(r'p)$, resp. (pq') et $(p'q)$) se coupent en u (resp. v, resp. w). Alors, u, v, w sont alignés.*

Démonstration. Si a, b sont deux points du plan on note (ab) à la fois la droite et une forme linéaire qui la définit. On définit $F = (qr')(rp')(pq')$ (c'est l'équation de degré 3 qui définit la réunion des trois droites) et $G = (q'r)(r'p)(p'q)$. On a $Z = \{p, q, r, p', q', r', u, v, w\}$. On appelle H une équation de Γ. Les 6 points p, q, r, p', q', r' sont sur $\Gamma = V(H)$, de sorte que les trois restants u, v, w sont sur $V(L)$, avec L de degré 1, c'est-à-dire une droite.

5.8 *Remarque.* Une méthode analogue permet de prouver le théorème de Pappus. Sur ces questions, on renvoie le lecteur à [GeoDP] Parties I et III.

5.2 Résultant et Nullstellensatz

Il s'agit du résultat suivant, forme affaiblie du théorème des zéros de Hilbert (les notations sont toujours celles de [IGADP]) :

5.9 Théorème. *Soit k un corps algébriquement clos, $P, Q \in k[X_1, \ldots, X_r]$. On suppose qu'on a $V(P) \subset V(Q)$. Alors, il existe $N \in \mathbf{N}$ tel que Q^N soit multiple de P.*

Démonstration. 1) On traite d'abord le cas où P est irréductible, donc non constant. On suppose par exemple que P est de degré > 0 par rapport à X_r et on appelle $R(X_1, \ldots, X_{r-1})$ le résultant de P, Q, vus comme polynômes en X_r. On a donc une relation $R = FP + GQ$, avec $d^\circ G < d^\circ P$ (par rapport à X_r). On va montrer que R est nul, ce qui donnera la conclusion. En effet, on a alors $FP = -GQ$ et, comme P est irréductible, il divise G ou Q, donc Q pour une raison de degré, et on a la conclusion.

Montrons donc que R est nul. Sinon, comme k est infini, il existe des scalaires $x_1, \ldots, x_{r-1} \in k$ tels que $R(x_1, \ldots, x_{r-1}) \neq 0$. Mais alors, comme k est algébriquement clos et P de degré > 0 en X_r, le polynôme $P(x_1, \ldots, x_{r-1}, X_r)$ admet une racine x_r. On a donc $P(x_1, \ldots, x_r) = 0$, donc aussi $Q(x_1, \ldots, x_r) = 0$ par hypothèse, donc $R(x_1, \ldots, x_{r-1}) = 0$, et c'est absurde.

2) On passe au cas général en écrivant $P = P_1^{\alpha_1} \cdots P_s^{\alpha_s}$ où les P_i sont irréductibles. Comme $V(P)$ est réunion des $V(P_i)$ on a $V(P_i) \subset V(Q)$ donc, par 1), P_i divise Q. Mais alors, si $N = \operatorname{Max} \alpha_i$, on voit que P divise Q^N.

5.3 Résultant et paramétrage

Le lecteur qui souhaiterait obtenir des précisions sur les courbes rationnelles (ou unicursales) pourra consulter [IGADP] Ch. IX.

5.3.1 Trouver l'équation cartésienne d'une courbe rationnelle

Donnons d'abord une définition un peu imprécise :

5.10 Définition. *Soit k un corps et soient $P(T)$ et $Q(T)$ deux fractions rationnelles à coefficients dans k. La courbe rationnelle associée à P, Q est l'image de l'application $\varphi : k \to k^2$ qui à t associe $(P(t), Q(t))$. On la note $\mathscr{C}(P, Q)$.*

5.11 *Remarques.* 1) Attention, l'application φ n'est pas définie en les pôles de P, Q.

2) Attention aussi, si l'on ne suppose rien sur P, Q l'image peut être finie. C'est le cas par exemple, si les fractions sont constantes.

5.12 Proposition. *La courbe $\mathscr{C}(P, Q)$ est incluse dans une courbe algébrique $V(R)$, avec $R \in k[X, Y]$.*

Démonstration. On écrit $P = \dfrac{P_1}{P_2}$ et $Q = \dfrac{Q_1}{Q_2}$ avec les polynômes P_i (resp. Q_j) premiers entre eux et on considère le résultant $R(X,Y)$ des polynômes de $k[X,Y][T] : P_1(T) - XP_2(T)$ et $Q_1(T) - YQ_2(T)$. Alors on a $\mathscr{C}(P,Q) \subset V(R)$. En effet, si (x,y) est sur $\mathscr{C}$, il existe t tel que $x = P(t)$ et $y = Q(t)$, ce qui signifie que les deux polynômes $P_1(T) - xP_2(T)$ et $Q_1(T) - yQ_2(T)$ ont une racine commune, donc que leur résultant $R(x,y)$ est nul (par le principe de spécialisation).

5.13 *Exemples.* La preuve de la proposition précédente donne un moyen de calculer l'équation de la courbe comme résultant. Dans tous les exemples ci-dessous, les calculs sont menés avec le logiciel *xcas*.

1) On suppose $P(t) = \dfrac{t}{1+t^3}$ et $Q(t) = \dfrac{t^2}{1+t^3}$. On trouve le folium de Descartes : $R(x,y) = x^3 + y^3 - xy$.

2) On suppose $P(t) = t^5 + 2t^4 - t^3 + t^2 - t + 1$ et $Q(t) = t^2 + 1$. On trouve la quintique :

$$R(x,y) = -y^5 + 11y^4 - 29y^3 - 4xy^2 + x^2 + 34y^2 + 6xy - 4x - 19y + 5.$$

3) On suppose $P(t) = \dfrac{t^2-1}{t^2+1}$ et $Q(t) = \dfrac{t(t^2-1)}{t^2+1}$. On trouve la cubique singulière à l'origine :

$$R(x,y) = 4(x^3 + xy^2 + x^2 - y^2).$$

4) On suppose $P(t) = \dfrac{t^2+1}{2t}$ et $Q(t) = \dfrac{2t-1}{t^2}$. On trouve la cubique (singulière en $(1,1)$) :

$$R(x,y) = 4x^2y - 4xy + y^2 - 4x - 2y + 5.$$

5) Pour des exemples avec des courbes de Bézier voir :
`http://www.math.u-psud.fr/~perrin/CAPES/geometrie/BezierDP.pdf`
sur ma page web.

5.3.2 Le cas algébriquement clos

Sur un corps quelconque, un paramétrage φ, s'il est à valeurs dans une courbe, n'est pas nécessairement surjectif. Par exemple, si $k = \mathbf{R}$, l'application définie sur $\mathbf{R} - \{\pm 1\}$ par $\varphi(t) = (\operatorname{ch} t, \operatorname{sh} t)$ est à valeurs dans l'hyperbole $V(X^2 - Y^2 - 1)$ mais n'atteint que la branche $x > 0$.

Sur un corps algébriquement clos, dans le cas d'un paramétrage polynomial, on a le résultat suivant :

5.14 Proposition. *On suppose k algébriquement clos. Soit $F \in k[X,Y]$ un polynôme irréductible non nul et soit $\varphi : k \to V(F)$ une application polynomiale non constante : $\varphi(t) = (P(t), Q(t))$ avec $P, Q \in k[T]$, non tous deux constants. Alors, φ est surjective.*

Démonstration. On a vu que (x, y) est dans $\mathscr{C}(P, Q) = \operatorname{Im}\varphi$ si et seulement si on a $R(x, y) = 0$ où R est le résultant de $P(T) - X$ et $Q(T) - Y$ par rapport à T. On a donc $V(R) \subset V(F)$. De plus, R ne peut être constant. En effet, s'il était nul on aurait $V(R) = k^2$, ce qui est absurde car $V(R)$ est contenu dans $V(F)$ et F est non nul. S'il était constant et non nul, $V(R)$ serait vide donc aussi $\operatorname{Im}\varphi$ et c'est absurde.

Il s'ensuit que $V(R)$ est infini et différent de k^2. Le petit Nullstellensatz 5.9 montre alors que R vise F^N, mais comme F est irréductible, cela montre que R n'a pas d'autre facteur premier que F, donc on a bien $V(R) = V(F)$.

5.4 L'exemple de Delahaye

Il s'agit d'expliquer le phénomène suivant que Jean-Paul Delahaye évoque dans son livre *Merveilleux nombres premiers*, Belin, 2000.

Soit n un entier. Les nombres $n^{17} + 9$ et $(n+1)^{17} + 9$ sont-ils toujours premiers entre eux ?

On montre que c'est vrai pour tout n jusqu'à des milliards de milliards, mais il y a un contre-exemple pour :

$$n = 8424432925592889329288197322308900672459420460792433,$$

le facteur commun étant :

$$p = 893658223791571665995096225335894563579345325693555 9.$$

L'explication est simple : on calcule le résultant des polynômes $P(x) = x^{17} + 9$ et $Q(x) = (x+1)^{17} + 9$ dans $\mathbf{Z}$. On trouve le nombre p. Cela montre que les deux polynômes ont un facteur commun dans $\mathbf{Z}/p\mathbf{Z}$. On cherche donc le *pgcd* de ces polynômes modulo p. On trouve le polynôme $x - n$. On a donc une racine commune n entre ces deux polynômes modulo p, ce qui signifie que p divise $n^{17} + 9$ et $(n+1)^{17} + 9$.

Il n'est pas difficile de bâtir d'autres exemples de ce style, par exemple $n^{19} + 6$ et $(n+1)^{19} + 6$: ils sont premiers entre eux jusqu'à

$$n = 372160549911586950840693700787968824987002420679664522043703 9,$$

mais pour cet entier admettent le facteur commun :

$$p = 529987588867054956554872480812165989490203291692575255926283 7.$$

5.15 *Remarque.* Bien entendu, l'intérêt de ces exemples est de montrer qu'en mathématiques, il ne suffit pas de vérifier un résultat pour de nombreuses valeurs pour qu'il soit vrai !

6 Discriminant

Dans cette section, on s'intéresse aux racines communes entre un polynôme et sa dérivée, c'est-à-dire à ses racines multiples, qui conduisent à la notion de discriminant. On a choisi de travailler exclusivement sur un corps. Pour les notions de théorie des corps et de théorie de Galois utilisées dans ce paragraphe, on renvoie à [DP] ou à [TER].

6.1 Notations. On désigne par K un corps de caractéristique différente de 2, par P un polynôme de degré $n > 0$ à coefficients dans K et par L son corps de décomposition $L = D_K(P)$. On suppose que le polynôme P est séparable c'est-à-dire qu'il admet n racines distinctes dans L, que l'on note $x_1, \ldots, x_n$. L'extension L/K est alors galoisienne et on note G son groupe de Galois, qui s'injecte dans le groupe symétrique $\mathfrak{S}_n$ par la formule : $g(x_i) = x_{\sigma_g(i)}$ (voir [TER] 3.7).

6.1 Définition et propriété caractéristique

6.2 Proposition-Définition. *On pose* $\delta = \prod_{1\leq i<j\leq n} (x_i - x_j)$ *et* $\Delta = \delta^2 = \prod_{1\leq i<j\leq n} (x_i - x_j)^2$. *Le nombre*[2] Δ *est appelé* **discriminant** *du polynôme* P. *Les nombres* δ *et* Δ *sont des éléments de* L^* *et on a la formule*

$$\Delta = (-1)^{n(n-1)/2} \prod_{i\neq j} (x_i - x_j).$$

Le signe $(-1)^{n(n-1)/2}$ *est égal à* 1 *si* $n \equiv 0, 1 \pmod 4$ *et à* -1 *sinon.*

Démonstration. Il suffit de compter les signes $-$, donc les couples (i, j) avec $i > j$, il y en a bien $n(n-1)/2$.

6.3 *Remarque.* Attention, certains auteurs prennent $\Delta = \prod_{i\neq j}(x_i - x_j)$ comme définition du discriminant, mais la proposition suivante montre que c'est mal adapté à la théorie de Galois.

2. On le note $\Delta(P)$ lorsqu'on veut préciser de quel polynôme il est le discriminant.

6.4 Proposition. *Soit g un élément de G et σ_g la permutation associée.*
1) On a les formules $g(\delta) = \epsilon(\sigma_g)\delta$ et $g(\Delta) = \Delta$.
2) Le discriminant Δ est dans K^* (et pas seulement dans L^*).
3) On a les équivalences :

$$\delta \in K^* \iff \Delta \in K^{*2} \iff G \subset \mathfrak{A}_n.$$

Démonstration. La formule avec δ résulte du comptage du nombre d'inversions[3] de σ_g et celle avec Δ est évidente. Le point 2) en résulte car K est le corps fixe de G, voir [TER] 4.12. Enfin, le point 3) résulte lui aussi de 1) : si G est formé de permutations paires, les éléments de G fixent δ et inversement.

6.5 *Exemple.* Calculons le discriminant du polynôme du second degré $ax^2 + bx + c$. Ses racines sont x_1 et x_2 et on a $\Delta = (x_1 - x_2)^2 = (x_1 + x_2)^2 - 4x_1x_2 = \big(-\frac{b}{a}\big)^2 - 4\frac{c}{a} = \frac{b^2 - 4ac}{a^2}$.

6.2 Calcul du discriminant

6.6 Notations. On reprend les notations précédentes mais on suppose de plus que P est unitaire :

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

On suppose que la caractéristique du corps ne divise pas n. On considère le polynôme dérivé $P'(x)$ et on note $y_1, \ldots, y_{n-1}$ ses racines, dans une extension convenable. On a donc :

$$P'(X) = nX^{n-1} + (n-1)a_{n-1}X^{n-2} + \cdots + a_1 = n\prod_{j=1}^{n-1}(X - y_j).$$

6.7 Théorème. *Soit Δ le discriminant de P. On a les formules :*

$$\Delta = (-1)^{n(n-1)/2}\prod_{i=1}^{n} P'(x_i), \tag{1}$$

$$\Delta = (-1)^{n(n-1)/2}\prod_{i,j}(x_i - y_j), \tag{2}$$

3. Voire de la définition de la signature que l'on peut donner par cette formule, vue dans l'anneau de polynômes $K[x_1, \ldots, x_n]$.

$$\Delta = (-1)^{n(n-1)/2} n^n \prod_{j=1}^{n-1} P(y_j), \tag{3}$$

$$\Delta = (-1)^{n(n-1)/2} R(P, P'). \tag{4}$$

Démonstration. On part de la formule $P(X) = \prod_{i=1}^{n}(X - x_i)$ que l'on dérive :

$$P'(X) = \sum_{i=1}^{n}(X - x_1) \cdots \widehat{(X - x_i)} \cdots (X - x_n)$$

où le chapeau signifie que le terme correspondant est omis. On calcule alors $P'(x_i)$. Tous les termes de la somme sont nuls sauf celui où l'on a omis x_i et on a donc, pour i fixé, $P'(x_i) = \prod_{j,j\neq i}(x_i - x_j)$. On en déduit la valeur du produit $\prod_{i=1}^{n} P'(x_i) = \prod_{i,j,\, j\neq i}(x_i - x_j)$ et la première formule vient de 6.2. En utilisant l'expression de P' en fonction de ses racines, on a $P'(x_i) = n\prod_{j=1}^{n-1}(x_i - y_j)$ d'où $\prod_{i=1}^{n} P'(x_i) = n^n \prod_{i,j}(x_i - y_j)$ et la seconde formule. Mais on a aussi $P(y_j) = \prod_{i=1}^{n}(y_j - x_i)$ et donc $\prod_{j=1}^{n-1} P(y_j) = \prod_{i,j}(y_j - x_i)$. Par rapport à l'expression précédente, chaque terme $x_i - y_j$ est changé de signe, ce qui fait $n(n-1)$ changements. Comme ce nombre est pair, le signe est le même et on a bien la troisième formule. Enfin, la quatrième vient de 4.4.

Ces formules permettent de calculer le discriminant du polynôme du troisième degré :

6.8 Proposition. *Le discriminant de $P(X) = X^3 + pX + q$ est $\Delta = -4p^3 - 27q^2$.*

Démonstration. On calcule $P'(X) = 3X^2 + p$ dont les racines sont $y_j = \pm\sqrt{-\dfrac{p}{3}}$, $j = 1, 2$, et on vérifie que le produit $P(y_1)P(y_2)$ vaut $A = \dfrac{27q^2 + 4p^3}{27}$. On a alors $\Delta = -27A$ et le résultat.

6.9 *Exercice.* Montrer que le discriminant de $P(X) = X^n + pX + q$ est donné par la formule :

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

6.10 *Exercice.* Montrer que le discriminant de l'équation complète du troisième degré $aX^3 + bX^2 + cX + d = 0$ est égal à :

$$-27a^3d^2 + 18a^2bcd - 4a^2c^3 - 4ab^3d + ab^2c^2.$$

6.3 Un exemple : le cas cyclotomique

Si n est un entier positif, premier à la caractéristique de K, on rappelle qu'on note Φ_n le polynôme cyclotomique, polynôme unitaire, dont les racines sont les racines n-ièmes primitives de l'unité, voir [DP] Ch. III.

6.11 Proposition. *1) Soit n un entier quelconque premier à la caractéristique de K. On a $\Delta(X^n - 1) = (-1)^{(n-1)(n+2)/2}\, n^n$.*

2) Soit n un nombre premier impair. On a $\Delta(\Phi_n) = (-1)^{n(n-1)/2} n^{n-2}$.

Démonstration. 1) On peut calculer avec la formule utilisant les $P'(x_i)$, mais ici, il est bien plus simple d'utiliser l'autre. Si l'on pose $P(X) = X^n - 1$ on a $P'(X) = nX^{n-1}$ et son unique racine est 0. On a donc $P(y_j) = -1$ pour tout j et la formule en découle (c'est d'ailleurs un cas particulier de $X^n + pX + q$, voir exercice ci-dessus).

2) Ici, on va utiliser la formule avec les $P'(x_i)$. On a $X^n - 1 = (X-1)\Phi_n(X)$ ce qui donne $\Phi_n(X) = X^{n-1} + \cdots + X + 1$ et aussi, en dérivant, $nX^{n-1} = (X-1)\Phi_n'(X) + \Phi_n(X)$ et, si on applique cela avec $X = \zeta^i$, ζ racine n-ième primitive et $i = 1, \ldots, n-1$, on trouve $n\zeta^{i(n-1)} = (\zeta^i - 1)\Phi_n'(\zeta^i)$. On a donc $\Phi_n'(\zeta^i) = \dfrac{n\zeta^{i(n-1)}}{\zeta^i - 1}\cdot$ Comme on a $\zeta^n = 1$, le numérateur est égal à ζ^{-i} et le produit de ces termes est le coefficient constant de Φ_n, soit 1, au signe $(-1)^{n-1}$ près. Les $\zeta^i - 1$, eux, sont les racines du polynôme $\Phi_n(X+1) = (X+1)^{n-1} + \cdots + (X+1) + 1 = X^{n-1} + \cdots + n$ et leur produit est donc $(-1)^{n-1}n$. En définitive, on a $\Delta(\Phi_n) = (-1)^{n(n-1)/2} n^{n-1} \prod_{i=1}^{n-1} \Phi_n'(\zeta^i) = (-1)^{n(n-1)/2} n^{n-2}$.

6.4 Comment trouver des extensions de Q de groupe de Galois $\mathfrak{A}_3$?

6.4.1 Introduction

Parmi les problèmes naturels en théorie de Galois, on peut citer le problème "direct" qui consiste à calculer le groupe de Galois d'une extension donnée (par exemple le corps de décomposition d'un polynôme à coefficients rationnels) mais aussi le problème "inverse" où il s'agit, un groupe fini étant donné, de trouver une extension de $\mathbf{Q}$ dont ce groupe est le groupe de Galois. Ce

problème n'est pas résolu en général et nous n'abordons ici que les deux premiers cas non triviaux : trouver une extension de groupe $\mathfrak{S}_3$ ou $\mathfrak{A}_3 \simeq \mathbf{Z}/3\mathbf{Z}$.

On sait que, si $P(X) = X^3+pX+q$ avec p, q rationnels, le groupe de Galois G du corps de décomposition de P sur $\mathbf{Q}$ s'injecte dans $\mathfrak{S}_3$. Précisément, si P est irréductible, G contient le groupe alterné $\mathfrak{A}_3$ et il lui est égal si et seulement si $\Delta(P) = -4p^3 - 27q^2$ est un carré de $\mathbf{Q}$. On va montrer qu'on peut construire une infinité d'exemples dans les deux cas.

6.4.2 Exemples de groupes de Galois $\mathfrak{S}_3$

C'est le cas générique, il suffit d'être sûr que le discriminant n'est pas un carré de $\mathbf{Q}$ et il y un moyen infaillible pour cela qui consiste à prendre $p > 0$. En effet, Δ est alors négatif, donc n'est pas un carré. Encore faut-il s'assurer de l'irréductibilité de P. C'est très facile, comme le montre la proposition suivante :

6.12 Proposition. *1) Soit l un nombre premier. Le polynôme $X^3 + lX + l$ est irréductible sur $\mathbf{Q}$ et son groupe de Galois est $\mathfrak{S}_3$.*

2) Soit n un entier impair positif. Le polynôme X^3+nX+1 est irréductible sur $\mathbf{Q}$ et son groupe de Galois est $\mathfrak{S}_3$.

3) Soit n un entier positif congru à 2 modulo 3. Le polynôme X^3+nX+1 est irréductible sur $\mathbf{Q}$ et son groupe de Galois est $\mathfrak{S}_3$.

Démonstration. Le point 1) vient du critère d'Eisenstein, voir [DP] Ch. III, Th. 3.2. Pour 2) il suffit de réduire modulo 2. On obtient $X^3 + X + 1$ qui est irréductible, et on conclut par [DP] III, 3.5. Le point 3) est analogue en raisonnant modulo 3. Inutile de dire qu'on peut ainsi multiplier les exemples.

6.4.3 Exemples de groupes de Galois $\mathfrak{A}_3$

Le cas de $\mathfrak{A}_3$ est un peu plus difficile. On a vu que, pour obtenir des équations de groupe $\mathfrak{A}_3$, il faut résoudre dans $\mathbf{Q}$ l'équation $-4p^3-27q^2 = r^2$. C'est l'objet de la proposition suivante :

6.13 Proposition. *Les solutions rationnelles de l'équation $-4p^3-27q^2 = r^2$ (notée $(*)$) sont données par les formules suivantes :*

$$p = -\frac{1}{4}(a^2 + 27b^2), \quad q = -\frac{b}{4}(a^2 + 27b^2) \quad r = -\frac{a}{4}(a^2 + 27b^2)$$

avec $a, b \in \mathbf{Q}$.

Démonstration. On note que si p est nul, q et r le sont aussi et qu'on obtient cette solution en prenant $a = b = 0$ dans les formules ci-dessus. On peut donc supposer $p \neq 0$. On divise la relation $(*)$ par p et on obtient $-4p - 27\big(\frac{q}{p}\big)^2 = \big(\frac{r}{p}\big)^2$. On obtient les formules annoncées en posant $a = r/p$ et $b = q/p$. Inversement, on vérifie que ces formules donnent des solutions de $(*)$.

6.14 *Remarques.* 1) Pour obtenir des solutions entières de $(*)$, il suffit de prendre a, b entiers de même parité.

2) Attention, on peut obtenir des polynômes à coefficients entiers à partir de a, b rationnels. Par exemple, avec $a = \frac{6}{7}$ et $b = \frac{8}{7}$ on trouve $p = -9$ et $q = -72$.

3) Pour avoir une équation dont le groupe de Galois est $\mathfrak{A}_3$, encore faut-il que P soit irréductible. Ce n'est pas toujours le cas, par exemple $a = -\frac{20}{7}$ et $b = -\frac{6}{7}$ donnent le polynôme $X^3 - 7X + 6 = (X-1)(X-2)(X+3)$.

4) Attention, même si l'on prend a, b entiers, le polynôme P obtenu peut être réductible. Par exemple, avec $a = 70$ et $b = 12$, on obtient $X^3 - 2197X + 26364 = (X-13)(X-39)(X+52)$. En revanche, avec a, b impairs on a le résultat suivant.

6.15 Proposition. *Soient a, b des entiers impairs. Alors, si l'on pose $p = -\frac{1}{4}(a^2 + 27b^2)$ et $q = -\frac{b}{4}(a^2 + 27b^2)$, le polynôme $P(X) = X^3 + pX + q$ est à coefficients entiers, irréductible et a pour groupe de Galois $\mathfrak{A}_3$.*

Démonstration. Comme a^2 et b^2 sont congrus à 1 modulo 4 il est clair que P est à coefficients entiers. Vu 6.13, son discriminant est un carré. Il reste à voir que P est irréductible. Pour cela, nous aurons besoin de deux lemmes. Le premier concerne les polynômes du troisième degré :

6.16 Lemme. *Soit $P(X)$ un polynôme de degré 3 unitaire à coefficients entiers, réductible sur $\mathbf{Q}$ et dont le discriminant est un carré dans $\mathbf{Q}$. Alors, P admet trois racines entières.*

Démonstration. Notons u, v, w les racines de P. Comme $\mathbf{Z}$ est intégralement clos, il suffit de montrer que les trois racines sont rationnelles (voir par exemple [S] ou [ST]). Comme P est réductible sur $\mathbf{Q}$, l'une des racines, disons u, est rationnelle. Par ailleurs, comme les coefficients de P sont entiers, $u+v+w$ et $uv+uw+vw$ sont entiers et on en déduit que $v+w = (u+v+w)-u$ est rationnel ainsi que $vw = (uv+uw+vw)-u(v+w)$. Comme conséquence, on voit que $(u-v)(u-w) = u^2 - (v+w)u + vw$ est aussi rationnel.

Mais, le discriminant Δ est égal à δ^2 avec $\delta = (u-v)(u-w)(v-w)$ et l'hypothèse indique que δ est rationnel, donc aussi $v - w = \delta/(u-v)(u-w)$. Comme $v + w$ et $v - w$ sont rationnels, v et w le sont et on a gagné.

6.17 *Remarque.* La théorie de Galois donne aussi le résultat. En effet, si v et w n'étaient pas rationnelles, le corps de décomposition serait de degré 2 et le groupe de Galois contiendrait la transposition qui échange v et w, donc ne serait pas contenu dans $\mathfrak{A}_3$.

Le second lemme est purement arithmétique :

6.18 Lemme. *Soient u, v deux entiers premiers entre eux. Alors $uv(u+v)$ et u^2+uv+v^2 sont premiers entre eux.*

Démonstration. On raisonne par l'absurde en supposant que $uv(u+v)$ et u^2+uv+v^2 ont un facteur premier commun p. En vertu du lemme d'Euclide, p divise u, v ou $u+v$. S'il divise u, comme il divise u^2+uv+v^2, il divise v^2, donc v et c'est absurde. De même s'il divise v. S'il divise $u+v$, il divise aussi $(u+v)^2 = (u^2+uv+v^2)+uv$, donc aussi uv, donc u ou v et l'on est ramené à l'un des cas précédents.

Revenons à la proposition. Si P n'est pas irréductible, le lemme 6.16 montre que les trois racines de P sont entières et de somme nulle, ce sont donc $u, v, -u-v$ avec $u, v \in \mathbf{Z}$.

On en déduit $p = -u^2 - uv - v^2$, $q = uv(u+v)$, donc $b = \dfrac{q}{p} = \dfrac{-uv(u+v)}{u^2+uv+v^2}$. Soit d le *pgcd* de u, v. On a donc $u = du'$ et $v = dv'$ avec u' et v' premiers entre eux et $b = \dfrac{-du'v'(u'+v')}{u'^2+u'v'+v'^2}$. En vertu du lemme 6.18 appliqué à u', v' et du théorème de Gauss, b ne peut être un entier [4] que si $u'^2+u'v'+v'^2$ divise d, de sorte que b s'écrit $b = u'v'(u'+v')b'$ où b' est un entier. Mais, le nombre $u'v'(u'+v')$ est toujours pair et cela contredit le fait que b est impair.

6.19 *Exemples.* On obtient ainsi une infinité d'exemples de polynômes à coefficients entiers avec une extension de groupe de Galois $\mathfrak{A}_3$.

1) Pour $b = 1$ et $a = 2k+1$, on a $p = q = -(k^2+k+7)$, $\Delta = ((2k+1)(k^2+k+7))^2$.

2) On a une variante de 1) en prenant $b = -1$ ce qui change q en $-q$.

3) Pour $b = 3$ et $a = 2k+1$ on trouve $p = -(k^2+k+61)$ et $q = 3p$.

6.20 *Remarque.* Dans tous les exemples précédents, les coefficients entiers p et q ne sont jamais premiers entre eux. En effet, p divise q (et il ne peut être égal à -1 car a^2+27b^2 est ≥ 27 ; le cas $b = 0$ ne donne pas un polynôme irréductible). On peut donc se demander s'il existe des polynômes X^3+pX+q

4. L'exemple 6.14.4 a d'ailleurs été construit ainsi avec $u' = 1, v' = 3$ et $d = 13$.

à coefficients entiers, irréductibles, de groupe de Galois $\mathfrak{A}_3$ avec p, q premiers entre eux. C'est le cas, mais il faut prendre a, b rationnels. Par exemple $a = 1/13$ et $b = 5/13$ donnent [5] $p = -1$, $q = -5$ avec $\Delta = 26^2$. Pour une solution [6] avec $p < -1$ on peut prendre $a = 3$ et $b = 1/3$ qui donne $P(X) = X^3 - 3X - 1$.

7 Références

[DP] PERRIN Daniel, *Cours d'algèbre*, Ellipses, 1996.

[IGADP] PERRIN Daniel, *Géométrie algébrique, une introduction*, Interéditions, Paris, 1995.

[GeoDP] PERRIN Daniel, *Géométrie projective et applications aux géométries euclidienne et non euclidiennes*
`http://www.math.u-psud.fr/~perrin/Livre_de_geometrie_projective.html`

[S] SAMUEL Pierre, *Théorie algébrique des nombres*, Hermann, Paris, 1967.

[ST] STEWART Ian & TALL David, *Algebraic Number Theory*, Chapmann-Hall, 1987.

[TER] PERRIN Daniel, *La loi de réciprocité quadratique*, rédaction de TER (disponible pour les collègues sur simple demande).

5. On obtient des solutions avec $p = -1$ en résolvant l'équation diophantienne $\alpha^2 + 27\beta^2 = 4d^2$.

6. Ici, il faut trouver α, β, d tels qu $\alpha^2 + 27\beta^2$ divise $4d^3$. Par exemple $9, 1, 3$ ou $20, 6, 7$, etc.