Théorème de Ruffini et Abel

On donne dans cette note une démonstration d'un résultat dû à Ruffini (1799) et Abel (1824) énoncant qu'on ne peut pas donner de formule générique pour exprimer les racines d'un polynôme quelconque de degré 5 à l'aide uniquement des opérations élémentaires d'addition, soustraction, multiplication, division et d'un nombre fini de racines $k^{\text{ième}}$ itérées. Ce résultat est à mettre en regard du fait que de telles formules existent pour les polynômes de degré $2, 3$ et 4. La démonstration que l'on présente est due à V.I. Arnold. Quelques notations sont nécessaires pour poser le décor et énoncer le résultat. Pour $z = re^{i\theta} \in \mathbb{C}\backslash\{0\}$, avec $\theta \in [0, 2\pi)$, et $k \in \mathbb{N}$, posons

$$z^{1/k} = r^{1/k} \exp\left(i\Big(\frac{\theta}{k} + \frac{2\pi p}{k}\Big)\right),$$

où $0 \leqslant p < k$ est entier. On utilisera de facon abusive la notation $z^{1/k}$ pour l'un de k nombres ci-dessus. On notera F une fonction d'un nombre fini de variables complexes qui s'exprime à l'aide des opérations élémentaires d'addition, soustraction, multiplication, division. Une telle fonction peut être à valeurs dans $\mathbb{C}^\ell$ pour un certain $\ell \geqslant 1$. Deux occurences du symbole F dans une même formule signalent des fonctions de ce type-là a priori différentes. En notant $c = (c_0, c_1, c_2)$ on a par exemple

$$\frac{-c_1 + (c_1^2 - 4c_2c_0)^{1/2}}{2c_2} = F\big(c, F^{1/2}(c)\big). \tag{1}$$

On reconnait ici la formule donnant les racines d'un polynôme de degré 2. On notera F_1 une formule du type

$$F_1(c) = F\big(c, F^{1/k_1}(c)\big), \tag{2}$$

pour un entier $k_1 \geqslant 1$ et, pour $n \geqslant 2$, on notera F_n une formule du type

$$F_n(c) = F\big(c, F_{n-1}^{1/k_n}(c)\big)$$

pour un entier $k_n \geqslant 1$; elle met en jeu au plus n racines $k^{\text{ième}}$ itérées et les opérations élémentaires sur ces quantités. En d'autres termes, une formule de type F_1 peut contenir au plus une racine $k^{\text{ième}}$ et les opérations élémentaires sur ces quantités, comme dans $c_0^{1/2}$ ou $c_0^{1/2}(c_1 + c_2)^{1/3}$. On a ici

$$c_0^{1/2}(c_1 + c_2)^{1/3} = \big\{c_0^2(c_1 + c_2)^3\big\}^{1/6}$$

de la forme $F^{1/6}(c)$. Une formule de type F_2 peut contenir au plus deux racines de type $k^{\text{ième}}$ itérées et les opérations élémentaires sur ces quantités,, comme dans $\big(c_0 + (c_1c_2)^{1/3}\big)^{1/2}$ ou $\big(c_0 + (c_1c_2)^{1/3}\big)^{1/2}\big(c_0 + (c_1c_2)^{1/4}\big)^{1/5}$, et une formule de type F_3 trois racines de type $k^{\text{ième}}$ itérées et les opérations élémentaires sur ces quantités, comme dans $\big(c_0 + \big(c_1c_2 + (c_0)^{1/4}\big)^{1/3}\big)^{1/2}$.

Une remarque importante et déjà claire sur la définition de l'un des nombres $z^{1/k}$. L'image d'un lacet par une application de type F_1 n'est pas forcément un lacet. Sur l'exemple de $z^{1/k}$, suivons continument les coordonnées polaire (θ_t, ρ_t) de $z(t)$ d'un chemin $z(t)$ de classe C^1 ainsi que celles de son image $\big(\frac{\theta(t)}{k} + \frac{2\pi p}{k}, \rho_t^{1/k}\big)$. Comparez le cas où le lacet $(z(t))_{0\leqslant t\leqslant 1}$ décrit un petit cercle au voisinage d'un point $z_0 \neq 0$ au cas où le lacet décrit un certain nombre de tours complets autour du point 0. Le point $z(1)^{1/k}$ est l'image de $z(0)^{1/k}$ par une rotation d'angle $\frac{1}{k}\int_0^1 \dot\theta_t dt$. A fortiori l'image d'un lacet par une application de type F_n peut ne pas être un lacet.

Théorème $-$ *Quel que soit $n \geqslant 1$ il n'existe pas de formule de la forme*

$$r = F_n(c),$$

donnant l'ensemble $r \in \mathbb{C}^5$ des racines d'un polynôme unitaire quelconque sur $\mathbb{C}$ de degré 5 en terme de ses coefficients $c = (c_0, \ldots, c_4) \in \mathbb{C}^5$.

Pour démontrer ce réesultat on prend comme point de départ le fait que chaque coefficient d'un polynôme est une fonction symétrique élémentaire des racines r de ce polynôme, ce qu'on

écrit

$$c = f_{\mathrm{sym}}(r).$$

Si donc $(r(t))_{0\leqslant t\leqslant 1}$ désigne un chemin continu dans $\mathbb{C}^5$ tel que $r(1)$ est une permutation non triviale de $r(0)$ alors

$$c(t) := f_{\mathrm{sym}}(r(t))$$

décrit un lacet dans $\mathbb{C}^5$, du fait que la fonction f_{sym} est invariante par permutation de ses arguments. On va voir que pour un polynôme de degré 5, de racines distinctes r, pour toute formule de type F_n on peut trouver un chemin $(r(t))_{0\leqslant t\leqslant 1}$ issu de r tel que $r(1)$ est une permutation non triviale de r, en particulier $r(1) \neq r(0)$, et tel que

$$F_n(c(1)) = F_n(c(0)).$$

Cela empêche d'avoir l'identité $r = F_n(c)$, qui impliquerait la contradiction $r(1) = r(0)$ alors que $r(1) \neq r(0)$. Allons-y graduellement pour faire émerger le raisonnement. Un point notation d'abord: on paramètre dans la suite tous nos chemins par l'intervalle $[0, 1]$. La paramétrisation des chemins n'a aucune importance ici. Pour un lacet γ on note γ^{-1} le lacet parcouru dans le sens inverse, et pour deux lacets γ, γ' tels que $\gamma(1) = \gamma'(0)$ on note $\gamma \star \gamma'$ la concaténation des chemins consistant à d'abord suivre γ puis γ'.

(a) Une évidence, d'abord, qui met déjà en jeu le mécanisme fondamental. On ne peut pas décrire l'ensemble $r \in \mathbb{C}^2$ des racines d'un polynôme unitaire sur $\mathbb{C}$ de degré 2, de coefficients $c = (c_0, c_1, c_2) \in \mathbb{C}^3$, à l'aide d'une formule du type F, c'est-à-dire comme une fraction rationnelle des coefficients. Notons (r_1, r_2) le couple ordonné des racines d'un polynôme de degré 2, *génériquement distinctes*, et rappelons qu'une formule de type F envoie un lacet sur un lacet. Si nous avions $r = F(c)$ et si $(r(t))_{0\leqslant t\leqslant 1}$ désignait un chemin continu allant de (r_1, r_2) à (r_2, r_1), le chemin $c(t) = f_{\mathrm{sym}}(r(t))$ serait un lacet, et on aurait la contradiction

$$(r_1, r_2) = F\big(c\big) = F\big(c(0)\big) = F\big(c(1)\big) = (r_2, r_1).$$

La formule (1) permet cependant d'écrire $r = F_1(c)$ à l'aide d'une racine carrée.

(b) Prenons maintenant un polynôme unitaire de degré 3, ayant *trois racines distinctes* (r_1, r_2, r_3). Notons $(\sigma_{12}(t))_{0\leqslant t\leqslant 1}$ la famille d'application de $\mathbb{C}^3$ dans lui-même définie pour tout $r' \in \mathbb{C}^3$ et $0 \leqslant t \leqslant 1$ par

$$\sigma_{12}(t)(r') = \big((1-t)r'_1 + tr'_2\,,\, tr'_1 + (1-t)r'_2\,,\, r'_3\big). \tag{3}$$

Définissons pareillement

$$\sigma_{23}(t)(r') = \big(r'_1\,,\, (1-t)r'_2 + tr'_3\,,\, tr'_3 + (1-t)r'_2\big). \tag{4}$$

Les chemins associés au point r

$$\gamma_{ij}(t)(r) := f_{\mathrm{sym}}\big(\sigma_{ij}(t)(r)\big)$$

dans l'espace des coefficients $\{c\}$ sont des lacets de $\mathbb{C}^3$ du fait que f_{sym} est invariante par permutation de ses arguments. Comme on l'a noté juste avant l'énoncé du théorème, et avec la notation (2) pour F_1, le point $F^{1/k}\big(\gamma_{12}(1)(r)\big)$ est obtenu à partir de $F^{1/k}\big(\gamma_{12}(0)(r)\big)$ par une rotation d'un certain angle, disons ϕ. Le point $F^{1/k}\big(\gamma_{12}^{-1}(1)(r)\big)$ est alors obtenu à partir de $F^{1/k}\big(\gamma_{12}(0)(r)\big)$ par une rotation d'angle $-\phi$. Pareillement, le point $F^{1/k}\big(\gamma_{12} \star \gamma_{23}(1)(r)\big)$ est obtenu à partir de $F^{1/k}\big(\gamma_{12} \star \gamma_{23}(0)(r)\big)$ par une rotation d'un certain angle, disons ψ, et $F^{1/k}\big(\gamma_{12}^{-1} \star \gamma_{23}(1)(r)\big)$ est obtenu à partir de $F^{1/k}\big(\gamma_{12} \star \gamma_{23}(0)(r)\big)$ par une rotation d'angle $-\psi$. Il s'ensuit que $F^{1/k}$ envoie le *lacet*

$$\gamma(t)(r) := \big[\gamma_{12}(t), \gamma_{23}(t)\big](r) := \Big(\gamma_{12} \star \gamma_{23} \star \gamma_{12}^{-1} \star \gamma_{23}^{-1}\Big)(t)(r)$$

sur un *lacet* : les quatre angles s'ajoutent et sont de somme nulle ! Une application de type F_1 envoie donc elle aussi un lacet sur un lacet. On parle de γ comme du *commutateur des chemins* γ_{12} *et* γ_{23}. Dans le même temps on peut paramétrer le chemin

$$\sigma(t) := \big[\sigma_{12}\,,\, \sigma_{23}\big](t) := \Big(\sigma_{12} \star \sigma_{23} \star \sigma_{12}^{-1} \star \sigma_{23}^{-1}\Big)(t)(r)$$

pour avoir l'identité

$$f_{\mathrm{sym}}\big(\sigma(t)(r)\big) = \gamma(t)(r).$$

Ce chemin satisfait

$$r = \sigma(0)(r) \neq \sigma(1)(r) = (r_3, r_1, r_2).$$

Il n'est donc pas possible que $r = F_1(c)$, sans quoi on aurait

$$\sigma(t)(r) = F_1\big((f_{\mathrm{sym}} \circ \sigma(t))(r)\big) = F_1\big(\gamma(t)(r)\big)$$

le long du chemin, et la contradiction

$$\begin{aligned}\sigma(0)(r) = F_1(c) &= F_1\Big(\gamma_{12} \star \gamma_{23} \star (\gamma_{12})^{-1} \star (\gamma_{23})^{-1}(0)(r)\Big)\\ &= F_1\Big(\gamma_{12} \star \gamma_{23} \star (\gamma_{12})^{-1} \star (\gamma_{23})^{-1}(1)(r)\Big) = \sigma(1)(r).\end{aligned}$$

On sait par contre qu'on peut écrire $r = F_2(c)$, à l'aide d'une racine $k^{\text{ième}}$ itérée.

(c) Prenons dans ce paragraphe un polynôme de degré 4 ayant quatre *racines distinctes* $(r_1, \dots, r_4)$. Notons pour tout $r' \in \mathbb{C}^4$ et $0 \leqslant t \leqslant 1$

$$\sigma_{12}(t)(r) = \big((1-t)r'_1 + tr'_2\,,\, tr'_1 + (1-t)r'_2\,,\, r'_3\,,\, r'_4\big)$$

et définissons pareillement l'application $\sigma_{ij}(t)$ pour $i \neq j$ dans $\{1, \dots, 4\}$. Pour i, j, k, ℓ tous distincts le chemin

$$\sigma(t) := \big[[\sigma_{ij}(t), \sigma_{jk}(t)]\,,\, [\sigma_{jk}(t), \sigma_{k\ell}(t)]\big]$$

envoie $r = \sigma(0)(r)$ sur

$$\sigma(1)(r) = \big(r_{\nu(1)}, r_{\nu(2)}, r_{\nu(3)}, r_{\nu(4)}\big)$$

où ν est la permutation de $\{1, \dots, 4\}$

$$\nu = \big[[(ij), (jk)], [(jk), (k\ell)]\big] = (i\ell)(jk).$$

Prenons par exemple $(i, j, k, \ell) = (1, 2, 3, 4)$. L'image par f_{sym} du chemin $\big(\sigma(t)(r)\big)_{0 \leqslant t \leqslant 1}$ dans l'espace $\mathbb{C}^4$ des coefficients est le lacet

$$\gamma := \big[[\gamma_{12}, \gamma_{23}], [\gamma_{23}, \gamma_{34}]\big],$$

où

$$\gamma_{ij}(t) = f_{\mathrm{sym}}\big(\sigma_{ij}(t)(r)\big).$$

Une fonction de type F_2 est de la forme

$$F_2(c) = F\Big(c, F^{1/k_1}\big(c, F^{1/k_2}(c)\big)\Big).$$

Chaque fonction $\Big([\gamma_{12}, \gamma_{23}], F^{1/k_2}\big([\gamma_{12}, \gamma_{23}]\big)\Big)$ et $\Big([\gamma_{23}, \gamma_{34}], F^{1/k_2}\big([\gamma_{23}, \gamma_{34}]\big)\Big)$ décrit un lacet d'après le paragraphe **(b)**. L'image par $F^{1/k_1}\big(c, F^{1/k_2}(c)\big)$ du commutateur des deux lacets $[\gamma_{12}, \gamma_{23}]$ et $[\gamma_{23}, \gamma_{34}]$ produit donc elle aussi un lacet il en va de même pour F_2. *Il n'est donc pas possible que* $r = F_2(c)$, sans quoi on aurait

$$\sigma(t)(r) = F_2\big((f_{\mathrm{sym}} \circ \sigma(t))(r)\big) = F_2\big(\gamma(t)\big)$$

le long du chemin, et la contradiction

$$(r_1, \dots, r_4) = \sigma(0)(r) = F_2(c) = F_2\big(\gamma(0)\big) = F_2\big(\gamma(1)\big) = \sigma(1)(r) = (r_3, r_4, r_1, r_2) \neq (r_1, \dots, r_4)$$

puisque toutes les coordonnées de r sont supposées distinctes. On sait par contre qu'on peut écrire $r = F_3(c)$, à l'aide d'une racine $k^{\text{ième}}$ itérée deux fois.

Démonstration du théorème – Un polynôme unitaire de degré 5 a génériquement cinq *racines distinctes* $(r_1, \dots, r_5)$. Définissons

$$\mathscr{F}_0 := \Big\{\text{3-cycles des permutations de } \{1, \dots, 5\}\Big\}$$

$$\mathscr{F}_n := \Big\{[a, b] = aba^{-1}b^{-1};\, a, b \in \mathscr{F}_{n-1}\Big\}$$

pour $n \geqslant 1$. La relation

$$[(ijk),(k\ell m)] = (jkm), \tag{5}$$

valable pour tout i,j,k,ℓ,m dans $\{1,\dots,5\}$, implique que chaque trois cycle appartient à tous les $\mathscr{F}_n$. Donnons-nous maintenant pour chaque $i \neq j \in \{1,\dots,5\}$ un chemin $\sigma_{ij}(t)$ d'applications de $\mathbb{C}^5$ échangeant r_i et r_j linéairement comme en (3) ou (4) et laissant fixes les autres coordonnées. La formule

$$\gamma_{ij} := f_{\mathrm{sym}}\big(\sigma_{ij}(t)(r)\big).$$

définit un lacet de l'espace $\{c\}$ des coefficients. Définissons aussi

$$\mathscr{C}_0 := \Big\{[\gamma_{ij},\gamma_{jk}]\,;\, i,j,k \text{ distincts}\Big\}$$
$$\mathscr{C}_n := \Big\{[\gamma_a,\gamma_b];\, \gamma_a,\gamma_b \in \mathscr{F}_{n-1}\Big\}, \quad n \geqslant 1.$$

Une récurrence élémentaire montre que, pour $n \geqslant 4$, l'image par une application de type F_{n-2} d'un lacet γ de $\mathscr{C}_n$ est un lacet. Le cas $n = 4$ est analysé dans le point **(c)**. Le point crucial est qu'un tel lacet est l'image par f_{sym} d'un chemin $\sigma(t)(r)$ de $\mathbb{C}^5$ obtenu en composant un certain nombre de σ_{ij}. On a donc en conséquence des identités

$$\big[(ij),(jk)\big] = (ijk)$$

et (5) l'égalité

$$\sigma(1)(r) = \big(r_{\nu(1)}, r_{\nu(2)}, r_{\nu(3)}, r_{\nu(4)}\big) \neq (r_1,\dots,r_5)$$

où ν est un trois cycle de $\{1,\dots,5\}$ pour un choix ad hoc des σ_{ij}. Il s'ensuit qu'il n'est pas possible que $r = F_{n-2}(c)$ sans quoi on aurait

$$\sigma(t)(r) = F_{n-2}\big((f_{\mathrm{sym}} \circ \sigma(t))(r)\big) = F_{n-2}\big(\gamma(t)\big)$$

le long du chemin, et la contradiction

$$(r_1,\dots,r_5) = \sigma(0)(r) = F_{n-2}(c) = F_{n-2}\big(\gamma(0)\big) = F_{n-2}\big(\gamma(1)\big) = \sigma(1)(r).$$

On obtient la conclusion du fait que $n \geqslant 4$ est arbitraire. ▷

References

[1] V.B. Alekseev, *Abel's theorem in problems and solutions.* Kluwer Academic Publishers, (2004).

[2] L. Goldmakher, *Arnold's elementary proof of the insolvability of the quintic.* https://web.williams.edu/Mathematics/lg5/394/ArnoldQuintic.pdf.

[3] P. Ramond, *Abel-Ruffini's Theorem: Complex but Not Complicated!* arXiv:2011.05162, (2020).