CaMUS **2** (2012), 1–8

Du théorème de Rolle à la théorie de Khovanskii

Christiane Rousseau

1. Introduction

Beaucoup de problèmes de modélisation se ramènent à solutionner un système d'équations et à compter le nombre de solutions possibles. Des sous-problèmes importants consistent à borner le nombre de solutions d'un système, ou encore à montrer que le nombre de solutions est fini. Dans cet article, notre matériel de base est le théorème de Rolle, et nous montrons comment Askold Khovanskii est parti de questions très simples pour construire une théorie très puissante et élégante : la théorie des « fewnomials ». Michel Rolle est un mathématicien français à cheval sur le 17e et le 18e siècle (1652-1719). Askold Khovankii est un mathématicien russe. Le congrès international se tient tous les 4 ans. Les conférenciers invités sont les mathématiciens dont les travaux de recherche représentent les développements les plus significatifs dans les années précédant le congrès. Pour sa théorie des fewnomials. Askold Khovanskii a été conférencier invité au congrès international des mathématiciens de 1983 à Varsovie.

Rappelons le théorème de Rolle qui est illustré à la figure 1.

Théorème 1.1. *Si une fonction* $f : [a,b] \to \mathbb{R}$ *est continue sur* $[a,b]$ *et dérivable sur* $]a,b[$, *et si* $f(a) = f(b)$, *alors il existe* $c \in]a,b[$ *tel que* $f'(c) = 0$.

Exemple 1.2. *Tout polynôme* $P(x)$ *à coefficients réels de degré* n *a au plus* n *racines rélles. Ceci se prouve facilement par induction. En effet, c'est vrai pour un polynôme de degré* 1. *Supposons maintenant que ce soit vrai pour tout*

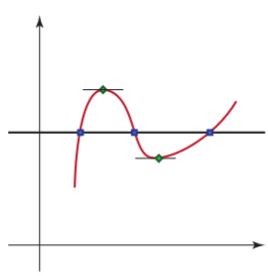

Figure 1. Le théorème de Rolle

polynôme de degré n, et soit $P(x)$ un polynôme de degré $n+1$. Supposons que $P(x)$ ait au moins $n+2$ racines réelles $x_1 < x_2, \cdots < x_{n+2}$. Alors, entre deux racines consécutives x_j and x_{j+1} il existe au moins une racine de la dérivée $P'(x)$. Ceci donne au moins $n+1$ racines pour P'. Contradiction, puisque P' est de degré n. Cette propriété de finitude du nombre de racines d'un polynôme $P(x)$ vient du caractère algébrique *de la fonction $P(x)$. Un fait remarquable est que le théorème de Rolle permet aussi de compter les racines multiples.*

Cependant, ce ne sont pas tous les polynômes de degré n qui ont de l'ordre de n racines. En effet, considérons le polynôme $x^n - 1$. Quelle que soit la taille de n, ce polynôme a au plus deux racines réelles.

Pourquoi ?

Réponse

- Parce qu'il a peu de monômes. Un théorème général résumant cette propriété est le théorème de Descartes que l'on décrira ci-dessous.
- Mais Khovanskii est allé plus loin. **Pourquoi un polynôme avec peu de monômes a-t-il peu de racines réelles ?** Nous allons voir sa réponse.

Commençons par une observation qui va nous préparer au théorème de Descartes.

PROPOSITION 1.3. *Un polynôme $P(x) = \sum_{i=1}^{m} a_i x^{n_i}$, où $n_1 < n_2 < \cdots < n_m$ et $a_i \neq 0$ pour tout i, a au plus $m-1$ racines réelles positives et au plus $m-1$ réelles négatives, en comptant les multiplicités. Il a donc au plus $2(m-1)+n_1$ racines réelles en comptant les multiplicités.*

DÉMONSTRATION. La preuve se fait par induction sur m. Si $m = 1$, on a un seul monôme qui ne s'annule qu'en $x = 0$ avec multiplicité n_1. Supposons que la proposition soit vraie pour un polynôme ayant m monômes, et considérons $P(x) = \sum_{i=1}^{m+1} a_i x^{n_i}$. Il suffit de s'occuper des racines positives et de remarquer que les racines négatives sont les racines positives du polynôme $P(-x)$. Alors x est une racine positive de $P(x)$ si et seulement si x est une racine positive de $Q(x) = \frac{P(x)}{x^{n_1}}$. On a $Q(x) = a_1 + \sum_{i=2}^{m+1} a_i x^{s_i}$, où $s_i = n_i - n_1 > 0$. Entre deux racines positives de $Q(x)$ il existe une racine positive de $Q'(x)$. Mais $Q'(x) = \sum_{i=2}^{m+1} s_i a_i x^{s_i - 1}$ a au plus $m-1$ racines positives par l'hypothèse d'induction. Donc, $Q(x)$ a au plus m racines positives. □

Déjà nous voyons une généralisation importante : dans la preuve de la proposition, lorsque nous avons déterminé une borne supérieure pour le nombre de racines positives, nous n'avons utilisé à aucun moment le fait que les n_i sont des entiers. La proposition reste vraie avec la même preuve si les n_i sont des nombres réels ! C'est le corollaire

COROLLAIRE 1.4. *Une fonction $f(x) = \sum_{i=1}^{m} a_i x^{\alpha_i}$, où $\alpha_1 < \alpha_2 < \cdots < \alpha_m$ et $a_i \neq 0$ pour tout i, $\alpha_i \in \mathbb{R}$, a au plus $m-1$ racines réelles positives en comptant les multiplicités.*

La fonction $f(x)$ du corollaire 1.4 n'est plus algébrique sur $]0, \infty[$, mais seulement analytique. Pourtant, elle a, comme un polynôme, un nombre fini de racines.

Pourquoi ?

Nous y reviendrons. Mais nous allons commencer par raffiner la proposition 1.3 pour aboutir au théorème de Descartes. En effet, pour qu'un polynôme $P(x)$ ait des racines positives, il faut que tous ses coefficients ne soient pas de même signe.

THÉORÈME 1.5. *(théorème de Descartes) Soit* $P(x) = \sum_{i=1}^m a_i x^{n_i}$, *où* $n_1 < n_2 < \cdots < n_m$ *et* $a_i \neq 0$ *pour tout* i, *un polynôme, et soit* r *le nombre de changements de signe dans la suite des coefficients* $a_1, \ldots, a_m$. *Alors* $P(x)$ *a au plus* r *racines réelles positives en comptant les multiplicités.*

DÉMONSTRATION. La démonstration suivante utilisant le théorème de Rolle a été donnée par Laguerre. Elle se fait par induction sur r. Le théorème est vrai pour $r = 0$. Supposons qu'il soit aussi vrai si on a r changements de signe et considérons un polynôme $P(x)$ avec $r+1$ changements de signe. Nous allons écrire $P(x) = \sum_{j=0}^{r+1} p_j(x)$, où $p_j(x)$ est une somme de monômes dont les coefficients sont de même signe, et les coefficients de p_j et p_{j+1} sont de signe opposé. Il est clair que x^α n'a pas de racine réelle positive pour tout nombre réel non nul α. Donc, le nombre de racines réelles positives de $P(x)$ est le même que le nombre de racines positives de $Q(x) = \frac{P(x)}{x^\alpha}$. Par le théorème de Rolle, ce nombre est au plus 1 plus le nombre de racines positives de $Q'(x)$. On a que

$$Q'(x) = \frac{\sum_{j=0}^{r+1} q_j(x)}{x^{\alpha+1}},$$

où $q_j(x) = p_j'(x)x - \alpha p_j(x)$. Le nombre de racines positives de $Q'(x)$ est bien sûr celui de $R(x) = \sum_{j=0}^{r+1} q_j(x)$. Pour montrer le théorème, il suffit de voir qu'on peut bien choisir le α de sorte que $R(x)$ n'ait que r changements de signe, et donc, au plus r racines. Regardons maintenant chacun des q_j. Supposons que $p_j(x) = \sum_{i=1}^{s_j} b_{ij} x^{k_{ij}}$ où $k_{1j} < k_{2j} < \cdots < k_{s_j j}$. Alors $q_j(x) = \sum_{i=1}^{s_j} b_{ij}(k_{ij} - \alpha)x^{k_{ij}}$. Il suffit donc de choisir α tel que, par exemple, $k_{s_0 0} < \alpha < k_{11}$, c'est-à-dire que α est compris entre le plus grand exposant de p_0 et le plus petit exposant de p_1. Alors, le signe des coefficients des q_j est le même que celui des coefficients des p_j pour $j \geq 1$, et l'inverse du signe des coefficients de p_0 quand $j = 0$. Donc, on a éliminé un changement de signe en passant de $P(x)$ à $Q(x)$, soit le changement de signe entre p_0 et p_1. Ainsi, de par l'hypothèse d'induction, $R(x)$ a au plus r changements de signe, et donc, r racines. Par suite, $P(x)$ a au plus $r + 1$ racines. □

Encore une fois, nous aurions pu remarquer que le théorème serait resté vrai si les n_i avaient été des nombres réels plutôt que des entiers.

Ces théorèmes se généralisent aux systèmes de n polynômes à n variables, mais nous nous contenterons de citer les résultats sans les prouver.

THÉORÈME 1.6. *(théorème de Bezout) Si* $P(x, y)$ *et* $Q(x, y)$ *sont deux polynômes à coefficients réels de degrés respectifs* m *et* n, *alors le système d'équations*

$$\begin{cases} P(x, y) = 0, \\ Q(x, y) = 0, \end{cases}$$

a au plus mn *solutions* (x_i, y_i) *dans* $\mathbb{R}^2$ *en comptant les multiplicités.*

Il y a moyen de donner une preuve de ce théorème en utilisant la théorie de l'élimination et en se ramenant au cas d'un polynôme à un variable. Ce qui nous intéresse plutôt, ce sont les généralisations qu'en a données Khovanskii dans le cas de systèmes où $P(x, y)$ et $Q(x, y)$ ont un petit nombre de monômes.

Théorème 1.7. *(Khovanskii) Le nombre de solutions réelles non dégénérées d'un système de* n *équations polynomiales à* n *inconnues ayant* q *monômes est inférieur ou égal à*

$$2^{q(q-1)/2}(n+1)^q.$$

Nous avons laissé en suspens deux grandes questions auxquelles Khovanskii a donné une réponse.

(1) Pourquoi les polynômes ou systèmes polynomiaux ayant peu de monômes ont peu de solutions réelles ?

(2) Pourquoi certaines fonctions analytiques ont des propriétés de finitude comme les fonctions algébriques ?

Ces questions ne sont que des cas particuliers de la grande théorie créée par Khovankii qui continue, depuis plus de 25 ans, d'être raffinée pour solutionner des problèmes de recherche de plus en plus complexes.

2. La théorie des fewnomials

La théorie des fewnomials explique pourquoi les polynômes ou systèmes polynomiaux ayant peu de monômes ont peu de solutions réelles. En effet, revenons sur l'exemple du polynôme $P(x) = x^n - 1$. Ce polynôme a n racines complexes qui sont les n racines de l'unité. Ces racines ont

– même module,

– des arguments uniformément répartis dans $[0, 2\pi]$.

La première propriété ne se généralisera pas pour un polynôme quelconque ayant un petit nombre de monômes, mais la deuxième restera valide sous une forme plus faible. Voici le théorème montré par Khovanskii. Il est un peu technique, mais il faut faire l'effort d'en comprendre l'idée.

Théorème 2.1. *On considère un système* $P(x_1, \ldots, x_n) = 0$*, où* $P = (P_1, \ldots, P_n)$ *et les* $P_j(x_1, \ldots, x_n)$ *sont des polynômes dans* $\mathbb{R}[x_1, \ldots, x_n]$*. Soit*

$$\mathbb{T}^n = \underbrace{\mathbb{S}^1 \times \cdots \times \mathbb{S}^1}_{n}$$

le tore de dimension n *qui est donné par le produit cartésien de* n *cercles. Sur* $\mathbb{T}^n$ *on considère des coordonnées angulaires* $(\theta_1, \ldots, \theta_n)$ *modulo* 2π*. Soit* Ω *une région de* $\mathbb{T}^n$*. À chaque solution de composantes non nulles* $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ *de* $P(x) = 0$*, on associe le vecteur* $\theta = (\theta_1, \ldots, \theta_n)$ *où* $\theta_j = \arg x_j$*. Soit* $N(P, \Omega)$ *le nombre de* θ *dans* Ω*. Alors, il existe des constantes* $\Pi = \Pi(\Omega, Newton)$ *et* $\phi(n, g)$*, où* Π *dépend de* Ω *et de l'enveloppe convexe des exposants des monômes du système (aussi appelé polygone de Newton), et* ϕ *dépend du nombre de monômes* g

du système $P(x) = 0$ et de la dimension n telles que

$$|N(P,\Omega) - N(P,\mathbb{T}^n)| \frac{\mathrm{vol}(\Omega)}{\mathrm{vol}(\mathbb{T}^n)} \leq \Pi\ \phi(g,n).$$

En pratique, le produit $\Pi\ \phi(g,n)$ est petit lorsque le système a peu de monômes, ce qui signifie que le quotient $\frac{N(P,\Omega)}{N(P\mathbb{T}^n)}$ est assez proche de $\frac{\mathrm{vol}(\Omega)}{\mathrm{vol}(\mathbb{T}^n)}$. Le concept de polygone de Newton est un concept technique qui ne sera pas décrit ici.

3. Les fonctions de Pfaff

Il existe une classe de fonctions analytiques qui ont de bonnes propriétés de finitude comme les fonctions algébriques. Ce sont les *fonctions de Pfaff*. La raison en est que ces fonctions ont une origine algébrique. Dans le cas des fonctions à une variable, ce sont des solutions d'équations différentielles algébriques :

(1)
$$\frac{dy}{dx} = \frac{P(x,y)}{Q(x,y)},$$

où $P(x,y)$ et $Q(x,y)$ sont des polynômes à coefficients réels. (En fait, toutes les solutions d'un tel système ne seront pas des fonctions de Pfaff, mais seulement les solutions *séparantes*. Nous discuterons cette distinction plus bas.)

Exemple 3.1. *(1) La fonction $y = x^\alpha$, où $\alpha \in \mathbb{R}$ est une fonction de Pfaff. En effet, elle est solution de l'équation différentielle*

$$\frac{dy}{dx} = \frac{\alpha y}{x}.$$

Rappel : ce sont des fonctions de cette forme qui ont été utilisées dans le corollaire 1.4.

(2) La fonction $y = e^{ax}$ est une fonction de Pfaff. En effet, elle est solution de l'équation différentielle

$$\frac{dy}{dx} = ay.$$

(3) La fonction $y = x \ln x$ est une fonction de Pfaff. En effet, elle est solution de l'équation différentielle

$$\frac{dy}{dx} = \frac{y}{x} + 1.$$

L'avantage de cette approche est que, lorsque nous considérons une équation différentielle, nous n'avons pas une fonction de Pfaff, mais toute une famille de fonctions de Pfaff qui sont solutions de la même équation différentielle. Cela a permis à Khovanskii de généraliser le théorème de Rolle sous la forme suivante.

Théorème 3.2. *(théorème de Rolle pour les systèmes dynamiques). On considère une région du plan, Ω, remplie de solutions d'une équation différentielle de la forme* (1)*, et une solution particulière (séparante) γ (voir figure* 2*). Soit $\mathcal{C}$ une courbe de classe C^1 dans Ω. Alors, entre deux points d'intersection consécutifs de γ et de $\mathcal{C}$ il existe sur $\mathcal{C}$ un point où $\mathcal{C}$ est tangente à une des solutions de l'équation différentielle* (1) *(un tel point est appelé un point de contact).*

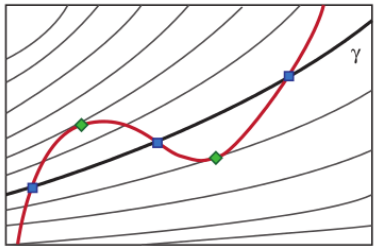

FIGURE 2. Le théorème de Rolle pour les systèmes dynamiques

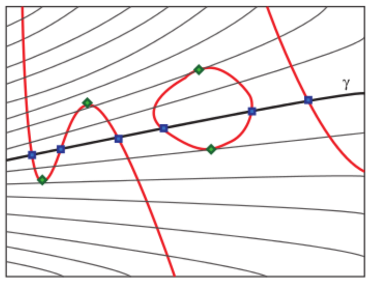

FIGURE 3. Points d'intersection et points de contact d'une solution séparante γ avec une courbe algébrique (en rouge)

La figure 2 illustre parfaitement la signification du théorème. Le théorème de Rolle que nous connaissons est un cas particulier de ce théorème. Il suffit de regarder l'équation différentielle $\frac{dy}{dx} = 0$, dont les solutions sont données par la famille des droites horizontales $y = C$.

DÉFINITION 3.3. *Une courbe algébrique de degré n est l'ensemble des points (x, y) du plan tels que $F(x, y) = 0$, où $F \in \mathbb{R}[x, y]$ est un polyôme en x, y à coefficients réels.*

COROLLAIRE 3.4. *Le nombre de points d'intersection isolés d'une courbe algébrique $F(x, y) = 0$ de degré m avec une solution séparante d'un système différentiel de la forme* (1) *de degré n est au plus $m(n+m)$. (Le degré d'une équation de Pfaff* (1) *est le maximum des degrés de P et Q.)*

DÉMONSTRATION. Sur chaque composante fermée, le nombre de points d'intersection est inférieur ou égal au nombre de points de contact. Sur chaque composante infinie, on a au moins un point de contact entre deux points d'intersection. Donc, le nombre de points d'intersection est inférieur ou égal au nombre de points de contact plus le nombre de composantes infinies (figure 3). Le nombre de points de contact est le nombre de solutions du système

$$\begin{cases} F_x Q + F_y P = 0, \\ F = 0, \end{cases}$$

puisque $F_xP + F_yQ = 0$ exprime que $\nabla F = (F_x, F_y)$ est perpendiculaire au champ de vecteurs $(P(x,y), Q(x,y))$. La première équation est polynomiale de degré $m - 1 + n$ et la deuxième de degré m. Donc, par le théorème de Bezout, le nombre de solutions est inférieur ou égal à $m(m + n - 1)$.

Il faut maintenant borner le nombre de composantes infinies. Les composantes infinies intersectent tout cercle $x^2 + y^2 = R$ en au moins deux points. Donc, le nombre de composantes infinies est inférieur ou égal à la moitié du nombre de points d'intersection de

$$\begin{cases} x^2 + y^2 = R, \\ F(x,y) = 0, \end{cases}$$

soit $\frac{1}{2}2m = m$. Le nombre de points d'intersection est donc inférieur ou égal à

$$m(m + n - 1) + m = m(m + n).$$

□

Regardons le miracle qui s'est produit : on a transformé la recherche du nombre de solutions d'un problème analytique en la recherche du nombre de solutions de deux problèmes algébriques. Donc, on a mis en évidence le caractère algébrique caché qui explique la finitude du nombre de solutions.

Définition 3.5. *On considère une équation de Pfaff* (1)*, dont on oriente les trajectoires selon la direction du champ de vecteurs tangents* $(P(x,y), Q(x,y))$*. Une solution (ou un ensemble fini de solutions),* γ*, est une* solution séparante *de l'équation de Pfaff si* γ *est la frontière orientée d'un domaine dans* $\mathbb{R}^2$*.*

La figure 4 donne des exemples de solutions non séparantes et séparantes. On voit que la condition que la solution est séparante est exactement ce qu'il faut pour que le théorème 3.2 fonctionne.

(a) spirale non séparante

(b) La réunion des deux droites extrêmes est non séparante.

(c) La réunion des deux droites extrêmes est séparante.

Figure 4. Exemples de solutions non séparantes et séparantes

Ce théorème se généralise à des systèmes dans $\mathbb{R}^n$ qui sont un mélange d'équations algébriques et d'équations pfaffiennes. De nombreuses variantes de la méthode existent que nous ne discuterons pas ici. Et les chercheurs qui utilisent la méthode créent les nouveaux raffinements que leurs systèmes d'équations exigent. Nous nous contenterons, pour terminer, de donner un exemple de théorème montré par Khovanskii.

Théorème 3.6. *Soit* $a_1, \dots, a_k \in \mathbb{R}^n$*. On considère un ensemble d'équations* $P_1(x_1, \dots, x_n) = \dots = P_n(x_1, \dots, x_n) = 0$*, où* P_j *est un polynôme de degré* d_j

dans les variables x_i *et les fonctions* $\exp(\langle a_s, x\rangle)$. *Alors, le nombre de racines non dégénérées de ce système dans* $\mathbb{R}^n$ *est au plus*

$$2^{\frac{k(k+1)}{2}}\left(\sum_{j=1^n}(d_j+1)\right)^k d_1\dots d_n.$$

4. Conclusion

Ce petit article veut illustrer qu'il n'y a pas de questions trop simples pour le scientifique à condition de garder les yeux ouverts, et qu'il n'est pas nécessaire d'avoir des connaissances très avancées pour faire des percées scientifiques significatives.

Références

[Kho82] A. Khovanskii, Cycles of dynamical systems on the plane and Rolle's theorem, *Funct. Anal. Appl.*, Sibirskii Math. Journal, **25** (1984), 502–506 (version originale en russe publiée en 1982).

[Kho] A. Khovanskii, Theory of fewnomials.

[Kho83] A. Khovanskii, Fewnomials and Pfaff manifolds, Proceedings of the International Congress of Mathematicians, 1983, Warszawa, 549–564.

CHRISTIANE ROUSSEAU, DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE MONTRÉAL
Courriel: rousseac@DMS.UMontreal.CA