Rolle's theorem: from a simple theorem to an extremely powerful tool

Christiane Rousseau, Université de Montréal

November 10, 2011

1 Introduction

How many solutions an equation or a system of equations have? Isn't that a fundamental question in mathematics? Indeed, so many modeling problems lead to solving systems of equations and count how many solutions they have. Important subproblems consist in bounding the number of solutions of a system, or showing the finiteness of the number of solutions of a system. This vignette is a variation on Rolle's theorem. The game was started by Askold Khovanskii. He started from very simple questions on the number of real roots of a polynomial with real coefficients. Looking to these with open eyes, he saw there the germ to build an extremely powerful theory, which he called the theory of fewnomials*. We invite you to follow the different steps of his approach.*

A small historical note. *Michel Rolle is a French mathematician from the end of the 17-th century (1652-1719). Askold Khovankii is a Russian mathematician. The International Congress of Mathematicians is held every four years. It acknowledges the most significant developments in mathematics in the years preceding the congress. For his work on fewnomials, Askold Khovankii was lecturer at the International Congress of Mathematicians in Warsaw in 1983.*

You all met Rolle's theorem in a calculus course:

Theorem 1 *If a function $f : [a, b] \to \mathbb{R}$ is continuous on [a,b] and differentiable on (a,b), and if $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$ (see Figure 1).*

Let us now explore some of its powerful applications.

Example 2 Any polynomial $P(x)$ with coefficients in $\mathbb{R}$ of degree n has at most n real roots. This can simply be proved by induction. Indeed, this is true for a polynomial of degree 1. Now, suppose it is true for all polynomial of degree n, and let $P(x)$ be a polynomial of degree $n+1$. By Rolle's theorem, between any two roots of $P(x)$ there is at least one root of $P'(x)$. Hence, the number of roots of $P(x)$ is at most 1 plus the number of roots of $P'(x)$. But, the degree of $P'(x)$ is n. Hence, P' has at most n roots, yielding that $P(x)$ has at most $n+1$ roots.

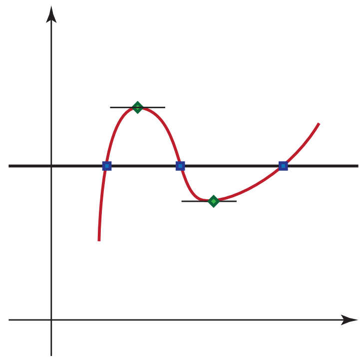


Figure 1: Rolle's theorem

However, not all polynomials of degree n have so many real roots. Indeed, consider the polynomial $x^n - 1$. No matter how large n, this polynomial has at most two real roots.

Why?

Answer.

- Because it has few monomials. This was already remarked by Descartes in the 17th century.
- But Khovanskii went further: why does a polynomial with few mononials has few real roots?

Let us explore these answers. But before, let us start with an observation that will lead us to Descartes's theorem.

Proposition 3 *A polynomial $P(x)$ with m monomials has at most $m-1$ positive real roots, and at most $m-1$ negative real roots, multiplicities taken into account. If the origin is a root of multiplicity k, then $P(x)$ has at most $2(m-1)+k$ real roots, multiplicities taken into account.*

Before going to the proof, let us consider an example:

Example 4 The polynomial

$$p(x) = x^{10^{100^{1000}}} - 1345x^3 + 1$$

has at most 2 positive roots, and at most 2 negative roots.

Proof of the proposition. The proof is by induction on m. If $m = 1$, there is a unique monomial which vanishes at $x = 0$ with multiplicity n_1. Let us suppose that the proposition is true for a polynomial with m monomials and let us consider a polynomial $P(x) = \sum_{i=1}^{m+1} a_i x^{n_i}$ with $m+1$ monomials. It suffices to consider the positive roots and to note that the negative roots are the positive roots of the polynomial $P(-x)$. Then x is a positive root of $P(x)$ if and only if x is a positive root of $Q(x) = \frac{P(x)}{x^{n_1}}$. We have $Q(x) = a_1 + \sum_{i=2}^{m+1} a_i x^{s_i}$, where $s_i = n_i - n_1 > 0$. Between two positive roots of $Q(x)$ there exists a positive root of $Q'(x)$. But $Q'(x) = \sum_{i=2}^{m+1} a_i x^{s_i - 1}$ has m monomials, and thus, it has at most $m-1$ positive roots by the induction hypothesis. Hence, $Q(x)$ has at most m positive roots. □

Already we see an important generalization: in the proof, to determine an upper bound for the number of positive roots we never used the fact that the n_i were integers. **Proposition 3 remains valid with the same proof if the n_i are real numbers!** This yields the corollary.

Corollary 5 *A function $f(x) = \sum_{i=1}^{m} a_i x^{\alpha_i}$, where $\alpha_1 < \alpha_2 < \cdots < \alpha_m$ and $a_i \neq 0$ for all i, $\alpha_i \in \mathbb{R}$, has at most $m-1$ positive real roots, multiplicities taken into account.*

The function $f(x)$ of Corollary 5 is no more algebraic on $]0, \infty[$, but only analytic. Nevertheless, it has, similarly to a polynomial, a finite number of roots.

Why?

We will come back to that later. Descartes's theorem is now just a refinement of Proposition 3 . It makes use of the sign of the coefficients of the polynomial. Indeed, if all coefficients of a polynomial $P(x)$ have the same sign, then $P(x)$ has no positive roots. Indeed, for $x > 0$, all monomials of $P(x)$ have the same sign.

Theorem 6 *(Descartes's theorem) Let $P(x) = \sum_{i=1}^{m} a_i x^{n_i}$, where $n_1 < n_2 < \cdots < n_m$ and $a_i \neq 0$ for all i, be a polynomial, and let r be the number of sign changes in the sequence of coefficients $a_1, \ldots, a_m$. Then $P(x)$ has at most r positive roots, multiplicities taken into account.*

If you prefer you can skip this proof on a first reading. Let us just mention that, again, the theorem remains true if the n_i are real numbers rather than integer numbers!

Proof of Descartes's theorem. The following proof was given by Laguerre, using induction on r. The theorem is true for $r = 0$. Let us uppose that it is true for a polynomial with r sign changes, and let us consider a polynomial $P(x)$ with $r+1$ sign changes. We will write $P(x) = \sum_{j=0}^{r+1} p_j(x)$, where $p_j(x)$ is a sum of monomials of same sign, and the coefficients of p_j and p_{j+1} are of opposite sign. It is clear that x^α has no real root for nonzero α. Hence, the number of positive roots of $P(x)$ is the same of the number of positive roots of $Q(x) = \frac{P(x)}{x^\alpha}$. By Rolle's theorem, this number is at most 1 plus the number of positive roots of $Q'(x)$. We have that

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

where $q_j(x) = p_j'(x)x - \alpha p_j(x)$. The number of positive roots of $Q'(x)$ is of course the same as the number of positive roots of $R(x) = \sum_{j=0}^{r+1} q_j(x)$. To prove the theorem, it suffices to see that we can choose α so that $R(x)$ has only r sign changes, and hence, at most r roots. Let us now look at each q_j. Suppose that $p_j(x) = \sum_{i=1}^{s_j} b_{ij} x^{k_{ij}}$ where $k_{1j} < k_{2j} < \cdots < k_{s_j j}$. Then, $q_j(x) = \sum_{i=1}^{s_j} b_{ij}(k_{ij} - \alpha) x^{k_{ij}}$. It suffices to choose α so that, for instance, $k_{s_0 0} < \alpha < k_{11}$, i.e. α is between the largest exponent of p_0 and the lower exponent of p_1. Then the sign of the coefficients of the q_j is the same as the sign of the coefficients of the p_j for $j \geq 1$, and the inverse of the signs of the coefficients of p_0 when $j = 0$. Hence, we eliminated one sign change when passing from $P(x)$ to $Q(x)$, namely the sign change between p_0 and p_1. Then, from the induction hypothesis, $R(x)$ has at most r sign changes, and hence, at most r roots. Thus, $P(x)$ has at most $r + 1$ roots. □

Example 2 can be generalized to systems of n polynomial equations in n variables. For $n = 2$ this is Bezout's theorem. We just state it.

Theorem 7 *(Bezout's theorem) If $P(x, y)$ and $Q(x, y)$ are two polynomials with real coefficients and respective degrees m and n, then the system of equations*

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

has at most mn solutions (x_i, y_i) in $\mathbb{R}^2$.

Khovanskii remarked that that there could be much less solutions when $P(x, y)$ and $Q(x, y)$ have a small number of monomials. For such polynomials, he introduced the colourful name of *fewnomials*. We state his result without proof.

Theorem 8 *(Khovanskii) The number of real (nondegenerate) solutions of a system of n polynomial equations with n unknowns and q monomials is less than or equal to*

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

We left aside two important questions to which Khovanskii gave an answer:

1. Why do polynomials or polynomial systems with few monomials have few real roots?
2. Why do certain analytic functions have finiteness properties similar to those of algebraic functions?

These questions are just particular cases of the important theory of Khovanskii which continues, for more than 25 years, to be refined so as to provide solutions to more and more complex research problems.

2 Why do polynomials with few monomials have few real roots?

This is explained by the theory of fewnomials. Indeed, let us come back to the example of the polynomial $P(x) = x^n - 1$. This polynomial has n complex roots which are the n-th roots of unity. These roots have

- same module,
- arguments uniformly distributed in $[0, 2\pi]$.

The first property cannot be generalized to an arbitrary polynomial with a small number of monomials, but the second property remains valid in a weaker form. It is even valid for systems of n polynomial equations in n unknowns. This result was presented by Khovanskii at the International Congress of Mathematicians in 1983.

The result of Khovanskii. We consider a system

$$\begin{aligned} P_1(x_1,\dots,x_n) &= 0,\\ &\vdots\\ P_n(x_1,\dots,x_n) &= 0. \end{aligned}$$

For each solution $(x_1,\dots,x_n)$ we consider the vector $(\theta_1,\dots,\theta_n)\in[0,2\pi]^n$, where θ_j is the argument of x_j. Then, the less monomials we have in the polynomials $P_1,\dots,P_n$, the more equidistributed are the points $(\theta_1,\dots,\theta_n)$ in the hypercube $[0,2\pi]$.

3 Analytic functions that behave like polynomials?

There exists a class of analytic functions which have good finiteness properties similarly to the algebraic functions. These are the *Pfaff functions*. The reason is that these functions have some algebraic origin. In the case of functions of one variable, they are solutions of algebraic differential equations, also called *Pfaff equations*:

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

where $P(x,y)$ et $Q(x,y)$ are polynomials with real coefficients. (In fact, not all solutions of Pfaff equations will be Pfaff functions, but only the *separating solutions*. We will explain this subtlety below.)

Example 9

1. The function $y=x^\alpha$, where $\alpha\in\mathbb{R}$, is a Pfaff function. Indeed, it is a solution of the differential equation

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

   Reminder: this is the form of the functions that have been used in Corollary 5.

2. The function $y=e^{ax}$ is a Pfaff function. Indeed, it is a solution of the differential equation

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

3. The function $y=x\ln x$ is a Pfaff function. Indeed, it is a solution of the differential equation

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

The advantage of this approach is that a Pfaff equation has *a whole family of solutions* and not a single one. Using this special feature, Khovanskii has generalized Rolle's theorem to the following:

Theorem 10 *(Rolle's theorem for dynamical systems). We consider a region of the plane, Ω, filled with solutions of a differential equation of the form* (1)*, and a particular (separating) solution γ (see Figure 2). Let $\mathcal{C}$ be a curve of class C^1 in Ω. Then, between two consecutive intersection points of γ with $\mathcal{C}$ there exists on $\mathcal{C}$ a point where $\mathcal{C}$ is tangent to one solution of the differential equation* (1) *(such a point is called a* contact point*.)*

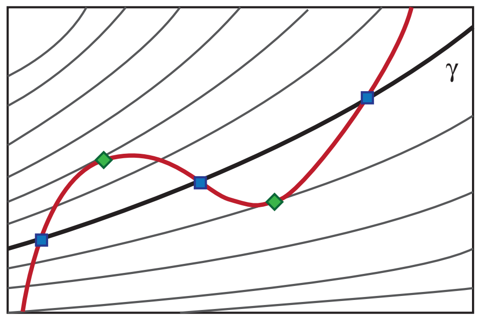


Figure 2: Rolle's theorem for dynamical systems.

Figure 2 highlights the meaning of the theorem. Rolle's theorem that we know is a particular case corresponding to the differential equation $\frac{dy}{dx} = 0$, whose solutions are the horizontal lines $y = C$.

Definition 11 An algebraic curve of degree n is the set of points (x, y) in the plane such that $F(x, y) = 0$, where F is a polynomial in x, y with real coefficients.

Corollary 12 *The number of isolated intersection points of an algebraic curve $F(x, y) = 0$ of degree m with a separating solution of a differential equation of the form* (1) *of degree n is at most $m(n + m)$. (The degree of a Pfaff equation* (1) *is the maximum of the degrees of P and Q.)*

Proof. On each closed component of $P(x, y) = 0$, the number of intersection points is less than or equal to the number of contact points (see Figure 3). On each infinite component, there is at least one contact point between two intersection points. Hence, the number of intersection points is less than or equal to the number of contact points plus the number of infinite components. Recall that the gradient of F, $\nabla F = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}\right)$, is orthogonal to the level curve $F(x, y) = 0$ of F. Hence, at a contact point, is it orthogonal to the vector field. Then, the number of contact points is the number of solutions of the algebraic(!) system

$$\begin{cases} \frac{\partial F}{\partial x} P + \frac{\partial F}{\partial y} Q = 0, \\ F = 0. \end{cases}$$

Both equations are polynomial with respective degrees $m - 1 + n$ and m. Hence, by Bezout's theorem, the number of solutions is less than or equal to $m(m + n - 1)$.

We now need to give a bound for the number of infinite components of $P(x, y) = 0$. These components intersect any circle $x^2 + y^2 = R$ in at least two points. Hence, the number

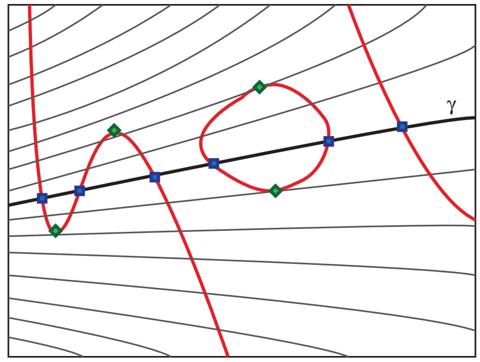


Figure 3: Intersection points and contact points between a separating solution γ and an algebraic curve (in red).

of infinite components is less than or equal to half the number of intersection points of $F(x,y)=0$ with a circle of sufficiently large radius R, i.e. the number of solutions of

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

which is bounded by $\frac{1}{2}2m=m$ by Bezout's theorem. Thus, the number of intersection points of the algebraic curve $P(x,y)=0$ with the separating solution is less than or equal to

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

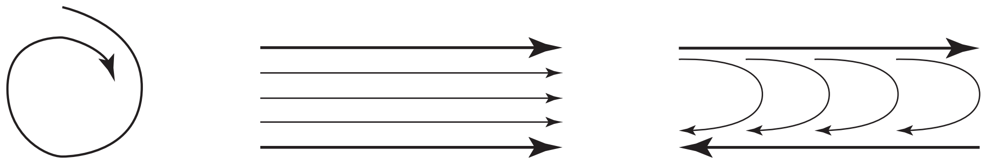
□

Using that the separating solution is one particular solution among a family of solutions of the Pfaff equation, we have reduced the problem of counting the number of solutions of a transcendental system to that of counting the number of solutions of an algebraic system! Hence, we could effectively use that $P(x,y)$ and $Q(x,y)$ are polynomials.

Definition 13 We consider a Pfaff equation (1), and we orient the solutions according to the direction of the tangent vector field $(P(x,y),Q(x,y))$. A solution (or a finite union of solutions), γ, is a *separating solution* of the Pfaff equation if γ is the oriented boundary of a domain in $\mathbb{R}^2$.

Figure 4 gives examples of non separating and separating solutions. We see that asking that the solution be separating is exactly the condition needed so that Theorem 10 be true.

This theorem can be generalized to systems in $\mathbb{R}^n$ which are a mixture of algebraic and Pfaffian equations. There exist many forms of the method which we will not discuss here. And researchers continue to prove new refinements that are needed for bounding the number of solutions of the systems coming from their modeling processes. A natural question is to ask how sharp are the bounds provided by Khovanskii's method. Unfortunately, they are not sharp and they may even by seven order of magnitudes larger than the real number of solutions of a system. So Khovanskii's method is more useful for proving finiteness results



(a) Non separating spirale.

(b) The union of the two extremal lines is non separating.

(c) The union of the two extremal lines is separating.

Figure 4: Examples of non separating and separating solutions

on the number of solutions and for providing explicit bounds (in opposition to proving the existence of a bound), than for deriving sharp bounds on the number of solutions.

4 Conclusion

This vignette highlights that no question is too simple for a mathematician, and that not all significant breakthroughs require advanced mathematics. We should not be afraid to look at the simplest things with the eyes of the kid and to ask "Why?"