

WINDING NUMBER AND THE NUMBER OF REAL ZEROS OF A FUNCTION

LIEVEN SMITS AND WILLEM KUYK

(Communicated by R. Daniel Mauldin)

ABSTRACT. The theorem in this paper shows that the number of real simple zeros of a function of the form $f(x) = q(x) + ax + b$, $x \in \mathbb{R}$, for not too wild $q(x)$ can be obtained counting the winding number of a closed plane curve about the point (a, b) .

1. INTRODUCTION AND PRELIMINARIES

Let q be a twice continuously differentiable function from \mathbb{R} to \mathbb{R} and put

$$f(x) = q(x) + ax + b, \quad x \in \mathbb{R},$$

where a and b are real constants. In this paper we shall first give a parametrization of the set of ordered pairs (a, b) such that f has a nonsimple zero, and then construct a geometric picture of the dependence of the number of zeros of f on a and b , provided that they are simple. We then determine the number of real zeros of f for a large class of functions q that includes all polynomial functions of degree greater than one.

Note that f has a nonsimple zero at $x \in \mathbb{R}$ if and only if

$$\begin{aligned} & \begin{cases} f(x) = 0 \\ f'(x) = 0 \end{cases} \\ & \iff \begin{cases} a = -q'(x) \\ b = xq'(x) - q(x). \end{cases} \end{aligned}$$

This parametrizes what we shall call the *catastrophe curve* $\text{cat}(q)$ of q in the (a, b) -plane. Note that curves of this type with $q(x)$ a polynomial occur in a natural way in the geometry of catastrophe types with one state parameter [1].

Lemma 1. $\text{cat}(q)$ is isotangent, i.e., the slope at a parameter point x is equal to $-x$.

Proof. One calculates

$$\begin{aligned} a'(x) &= -q''(x), \\ b'(x) &= q'(x) + xq''(x) - q'(x) = -xa'(x). \quad \square \end{aligned}$$

Received by the editors June 27, 1990.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 26C10; Secondary 12D10.

Both authors acknowledge the support of the Belgian National Science Foundation NFWO.

From now on we suppose that the function q satisfies the following axioms 1 and 3, and either axiom 2a or 2b:

- (1) $\lim_{x \rightarrow +\infty} q'(x) = +\infty$,
- (2a) $\lim_{x \rightarrow -\infty} q'(x) = +\infty$ (called the “odd case”),
- (2b) $\lim_{x \rightarrow -\infty} q'(x) = -\infty$ (called the “even case”),
- (3) q'' has all its zeros in a compact interval.

Remark. Note that these axioms are satisfied if q is a monic polynomial of degree at least two, and in that case the terms odd/even refer precisely to the degree of the polynomial.

Lemma 2. *The catastrophe curve $\text{cat}(q)$ tends to infinity in specific quadrants, more precisely:*

$$\begin{aligned} - \lim_{x \rightarrow +\infty} a(x) &= + \lim_{x \rightarrow +\infty} b(x) = +\infty, \\ \lim_{x \rightarrow -\infty} a(x) &= \lim_{x \rightarrow -\infty} b(x) = \pm\infty \text{ (even/odd)}. \end{aligned}$$

Proof. The statements for $a(\pm\infty)$ follow from $a(x) = -q'(x)$ and from the axioms. To compute $b(+\infty)$, choose $\alpha > 1$ such that $a'(x) < 0$ for $x \geq \alpha$. Next choose $\beta > \alpha$ such that $a(x) < a(\alpha) + b(\alpha)$ for $x \geq \beta$, which is possible on account of $a(+\infty) = -\infty$. By Cauchy’s mean value theorem, for all $x > \beta$, there exists a $\gamma \in (\beta, x)$ such that

$$a'(\gamma)(b(x) - b(\beta)) = b'(\gamma)(a(x) - a(\beta)).$$

As $\text{cat}(q)$ is isotangent, and since $a(x) - a(\beta) < 0$ there exists for all $x > \beta$ a $\gamma > \beta$ such that

$$\begin{aligned} b(x) &= b(\beta) - \gamma(a(x) - a(\beta)) \\ &\geq b(\beta) - \beta(a(x) - a(\beta)) \rightarrow +\infty \text{ as } x \rightarrow +\infty. \end{aligned}$$

Analogous proofs work for $b(-\infty)$ in the even and odd cases. \square

Lemma 3. *There exists an $R_0 > 0$ such that, if $R \geq R_0$, the circle $C(0, R)$ with radius R centered at the origin has two points in common with the image of $\text{cat}(q)$.*

Proof. We shall prove that, ultimately, the branches of $\text{cat}(q)$ go to infinity in a monotonic way. Consider the square norm

$$s(x) = a(x)^2 + b(x)^2.$$

First note that $s \geq a^2 = (q')^2 \rightarrow +\infty$. It follows that, for large radius R , there are at least two values of the parameter x where $\text{cat}(q)$ intersects $C(0, R)$.

Next, its derivative is

$$s' = 2aa' + 2bb' = 2q''(q' + x^2q' - xq).$$

By the axioms, q' has constant sign outside a compact interval. But so does $x^2q' - xq$, and the sign is the same:

$$xq'(x) - q(x) = q(0) + \int_{t=0}^x (q'(x) - q'(t))dt \rightarrow \pm\infty,$$

the sign depending on the cases odd/even and $x \rightarrow \pm\infty$. In fact, one verifies that s' is negative on a left half line, and positive on a right half line (with possibly several sign changes in between).

It follows that large circles intersect $\text{cat}(q)$ at exactly two values of the parameter. By Lemma 2, these values yield distinct points. \square

2. NOTATIONS AND THE CONCEPT OF LIMITING WINDING NUMBER

In the following we shall identify the topological one-torus, i.e. the interval $[0, 2\pi]$ with the end points identified, with the collection of all half lines in the plane with a fixed starting point. An *anticlockwise* motion in this torus (with respect to a fixed coordinate system, say XY) is a differentiable nonwinding (i.e. injective) path which encounters the unit vectors on the X and Y axes in the order $(e_X, e_Y, -e_X, -e_Y)$. A *clockwise* motion has the opposite orientation. An interval $[a, b]$ on the torus, for $a \neq b$, is chosen such that running through it is an anticlockwise motion. With this convention, we choose the range of the Arctg function to be $(-\pi/2, \pi/2) = (3\pi/2, \pi/2)$.

As before, let $C(0, R)$ be the circle in the (a, b) -plane with radius $R > 0$ centered at the origin. For any half line L in the plane, we denote by \tilde{L} the complementary half line with the same support, i.e., $L \cup \tilde{L}$ is a line.

By an *increasing* (resp. *decreasing*) zero of a real function we mean a value in its domain \mathbb{R} which is mapped to 0, together with a neighborhood of that value on which the function is strictly increasing (resp. decreasing). Classically, a simple zero of a continuously differentiable function is necessarily either an increasing or a decreasing one.

Definition. Let q, c, d and R_0 be as in the previous lemmas. The *limiting winding number* $W(q, c, d)$ for a point (c, d) in the plane not belonging to the image of $\text{cat}(q)$, is the winding number about (c, d) of a closed oriented curve S in $\mathbb{R}^2 \setminus \{(c, d)\}$, defined by the following construction.

Take an arbitrary $R > \max(R_0, \sqrt{c^2 + d^2})$ and let $x_0 < x_1$ be the parameter values of $\text{cat}(q)$ producing the intersection points of Lemma 3. The curve S is constructed by glueing together the following two pieces:

- $\text{cat}(q)$ from x_0 to x_1 ,
- a segment of $C(0, R)$, oriented anticlockwise from $(a(x_1), b(x_1))$ to $(a(x_0), b(x_0))$.

Note that this definition is independent of the particular choice of $R > \max(R_0, \sqrt{c^2 + d^2})$, since changing R is a homotopy of curves in $\mathbb{R}^2 \setminus \{(c, d)\}$ (one may construct the homotopy explicitly using a reparametrization of some patches by the parameter s of Lemma 3).

3. THE WINDING NUMBER THEOREM

Theorem. *In the odd case, the number of real zeros of*

$$f(x) = q(x) + ax + b$$

for $(a, b) \notin \text{cat}(q)$, is equal to $1 + 2W(q, a, b)$. In the even case, it is $2W(q, a, b)$.

Remark. The proof is rather technical. Figure 1 shows separately what is going on in the odd and even cases. The theorem concerns the closed bold curve that winds around the point (a, b) . The drawings of these curves are generic, in

the sense that no elements special to any particular case of $\text{cat}(q)$ are used. For the proof of the theorem it is not necessary to draw in each case that part of $\text{cat}(q)$ that lies inside the circle and that contains the cusps of $\text{cat}(q)$. Whence the incomplete drawings of the bold curves. The reader may wish to convince himself by drawing the curves $\text{cat}(q)$ for the special cases

$$q(x) = x^3 \quad (\text{odd case}), \quad q(x) = x^4 - 2x^2 \quad (\text{even case}).$$

Proof. Since $(a, b) \notin \text{cat}(q)$, all zeros are simple. We shall prove that the number of decreasing zeros of $f(x)$ is equal to $W(q, a, b)$. Since, in the odd case,

$$\lim_{x \rightarrow \pm\infty} q'(x) = +\infty,$$

we have

$$\lim_{x \rightarrow +\infty} f(x) = - \lim_{x \rightarrow -\infty} f(x) = +\infty$$

(with obvious analogues in the even case), and the theorem will be established.

Denoting a point (a, b) on $\text{cat}(q)$ by capital letters, observe first that, for all $x \in \mathbb{R}$, the tangent to $\text{cat}(q)$ at $(A(x), B(x))$ has the following equation in the (a, b) -plane

$$(1) \quad T_x \longleftrightarrow x \cdot a + b + q(x) = 0.$$

At the point $(a, b) = (A(x), B(x))$ of this line, we have

$$(2) \quad a + q'(x) = 0.$$

Furthermore, this line T_x has slope $-x \neq \infty$, so it cannot be parallel to the axis $b = 0$. We conclude that the quantity $\partial f / \partial x = a + q'(x)$, for the same fixed x , is negative for (a, b) on the open half line S_x of T_x which starts at $(A(x), B(x))$ and points into the negative a direction. Thus the real number x is a decreasing zero of $q(x) + ax + b$ if and only if (a, b) lies on the half line S_x .

The proof of the theorem proceeds in constructing a new closed curve $P = (p_1(\theta), p_2(\theta))$ in $\mathbb{R}^2 \setminus \{(0, 0)\}$ parametrized by $\theta \in [0, 2\pi]$, for which the crossings of the negative p_1 -axis in $\mathbb{R}^2 \setminus \{(0, 0)\}$ are in one-to-one correspondence with the decreasing zeros of $x \mapsto f(x) = q(x) + ax + b$ (a and b fixed).

To that end, suppose $R \geq R_0$ is so large that $C(0, R)$ contains all the zeros of q'' , which includes all the cusps of $\text{cat}(q)$. Then the two components of $\text{cat}(q)$ outside $C(0, R)$ are convex curves. It follows that each of them has at most one tangent line which passes through (a, b) . By making R larger if necessary, we ensure that all the lines tangent to $\text{cat}(q)$ which pass through (a, b) have their tangent points inside $C(0, R)$.

Let the two intersection points of $C(0, R)$ and $\text{cat}(q)$ correspond to the parameter values $x_0 < x_1$. We then have a curve S as described in §2. It lies in the nature of our problem that we construct P by letting an XY frame move about a fixed point, namely $(a, b) = (p_1(\theta), p_2(\theta))$; rather than letting a point move in \mathbb{R}^2 with respect to a fixed XY frame. In fact, we let the origin of an orthonormal XY frame slide over S in such a manner that, except at the points x_0 and x_1 , the X -axis is tangent to S . This makes for a parametrization of the form

$$\begin{aligned} p_1(\theta) &= (a - \alpha(\theta)) \cos \phi(\theta) - (b - \beta(\theta)) \sin \phi(\theta), \\ p_2(\theta) &= (a - \alpha(\theta)) \sin \phi(\theta) + (b - \beta(\theta)) \cos \phi(\theta), \end{aligned}$$

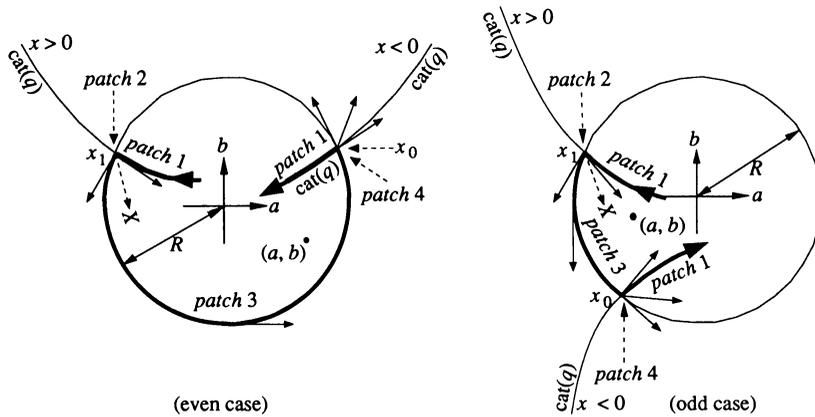


FIGURE 1

where the parameters $\phi(\theta)$, $\alpha(\theta)$ and $\beta(\theta)$ are yet to be suitably defined. Since S consists of four patches (see Figure 1), we shall describe how the positive X -axis will change direction on each of the four patches in terms of these three parameters (patches 2 and 4 are the points x_1 and x_0 respectively, where the X -axis rotates only). We offer two descriptions of these orientations, viz. an informal “kinematical” one and a formal one. The reader could skip any of the two if he so wishes.

Kinematical description of the patches:

- *patch 1.* For $\theta \in [\text{Arctg } x_0, \text{Arctg } x_1]$, the positive X -axis is the half line \tilde{S}_x that is tangent at $x = \text{tg } \theta$ to $\text{cat}(q)$, oriented towards positive a . Its starting point is $(A(\text{tg } \theta), B(\text{tg } \theta))$ on $\text{cat}(q)$.
- *patch 2.* For $\theta \in [\text{Arctg } x_1, \pi/2]$, the origin of the X -axis is fixed on $\text{cat}(q)$ at $x = x_1$; this is a point in the second quadrant of the (a, b) -plane. The direction of the X -axis moves clockwise, starting from the tangent to $\text{cat}(q)$ and ending at the anticlockwise tangent to $C(0, R)$.
- *patch 3.* For $\theta \in [\pi/2, 3\pi/2]$, the origin of the X -axis is a point on $C(0, R)$ moving anticlockwise from $\text{cat}(q)(x_1)$ to $\text{cat}(q)(x_0)$. The direction of the positive X -axis is the anticlockwise tangent to $C(0, R)$.
- *patch 4.* For $\theta \in [3\pi/2, \text{Arctg } x_0]$, the origin of the X -axis is fixed on $\text{cat}(q)$ at $x = x_0$; this is in the odd case (resp. even case) a point in the third (resp. first) quadrant of the (a, b) -plane. The direction of the positive X -axis moves, starting from the anticlockwise tangent to $C(0, R)$, in anticlockwise (resp. clockwise) fashion and, in both cases, ends up at the “positive a ” tangent to $\text{cat}(q)$.

Formal description of the patches:

In order to give the explicit values of the parameters $\phi(\theta)$, $\alpha(\theta)$ and $\beta(\theta)$ for the different patches, we introduce six auxiliary angles and agree for convenience that in the (a, b) -plane their clockwise fashion is counted positive.

They are given by the formulas

$$\begin{aligned} \operatorname{tg} \phi_0 &= x_0, & -\pi/2 < \phi_0 < 0, \\ \operatorname{tg} \phi_1 &= x_1, & 0 < \phi_1 < \pi/2, \\ R \cos \psi_0 &= A(x_0), & R \cos \psi_1 &= A(x_1), & \begin{cases} -\pi < \psi_0 < -\pi/2 & (\text{odd}), \\ 0 < \psi_0 < \pi/2 & (\text{even}), \\ \pi/2 < \psi_1 < \pi. \end{cases} \\ R \sin \psi_0 &= B(x_0), & R \sin \psi_1 &= B(x_1), \\ \phi'_0 &= -\pi/2 - \psi_0, & \begin{cases} 0 < \phi'_0 < \pi/2 & (\text{odd}), \\ -\pi < \phi'_0 < -\pi/2 & (\text{even}), \\ \pi/2 < \phi'_1 < \pi. \end{cases} \\ \phi'_1 &= 3\pi/2 - \psi_1, \end{aligned}$$

For the different patches the values are as follows:

- patch 1. ($\phi_0 \leq \theta \leq \phi_1$)

$$\phi(\theta) = \theta, \quad \alpha(\theta) = A(\operatorname{tg} \theta), \quad \beta(\theta) = B(\operatorname{tg} \theta).$$

- patch 2. ($\phi_1 \leq \theta \leq \pi/2$)

$$\begin{aligned} \phi(\theta) &= \phi_1 + \frac{\theta - \phi_1}{\pi/2 - \phi_1}(\phi'_1 - \phi_1), \\ \alpha(\theta) &= A(x_1), \quad \beta(\theta) = B(x_1). \end{aligned}$$

- patch 3. ($\pi/2 \leq \theta \leq 3\pi/2$)

$$\begin{aligned} \phi(\theta) &= \phi'_1 + \frac{\theta - \pi/2}{\pi}(\phi'_0 - \phi'_1), \\ \alpha(\theta) &= R \cos((3\pi/2 - \phi(\theta))), \quad \beta(\theta) = R \sin((3\pi/2 - \phi(\theta))). \end{aligned}$$

- patch 4. ($-\pi/2 \leq \theta \leq \phi_0$)

$$\begin{aligned} \phi(\theta) &= \phi'_0 + \frac{\theta + \pi/2}{\phi_0 + \pi/2}(\phi_0 - \phi'_0), \\ \alpha(\theta) &= A(x_0), \quad \beta(\theta) = B(x_0). \end{aligned}$$

Note that both the origin of the X -axis and its direction vary continuously as θ runs through the unit circle $[0, 2\pi]$. We claim that the *direction* of the X -axis has zero winding in the (a, b) -plane. In fact, it does not even cover the unit circle; e.g. it never has the direction of the negative a -axis. This implies that the winding number of the *origin* of the X -axis around (a, b) is equal to the winding number of \mathbf{S} around $(0, 0)$ in the (p_1, p_2) -plane.

We now prove that this number is also equal to the number of decreasing zeros of f . First note that by the choice of R , the negative X -axis never contains the point (a, b) in patches two, three and four (for the even case in patch four, look at the construction of R). It follows that \mathbf{S} does not cross the negative X -axis for $\theta \in [\operatorname{Arctg} x_1, \operatorname{Arctg} x_0]$. Next, for $\theta \in [\operatorname{Arctg} x_0, \operatorname{Arctg} x_1]$, the positive X -axis equals S_x , so that (cf. Equation (1)) crossings of the negative p_1 -axis correspond to $\operatorname{tg} \theta$ being a decreasing zero of f , while crossings of the positive p_1 -axis correspond to $\operatorname{tg} \theta$ being an increasing zero.

Hence, for $\theta \in [\operatorname{Arctg} x_0, \operatorname{Arctg} x_1]$, crossings of the p_1 -axis are alternating on its positive and negative halves, i.e., they correspond to increasing and

decreasing zeros, respectively. Therefore the number of each is the winding number of S indeed. This establishes the theorem. \square

REFERENCES

1. W. Kuyk and L. Smits, *On the geometries of the rational unfoldings of X^k* , Acta Appl. Math. (to appear).

UNIVERSITY OF ANTWERP (UIA), DEPARTEMENT WISKUNDE, UNIVERSITEITSPLEIN 1, B-2610 WILRIJK, BELGIUM

UNIVERSITY OF ANTWERP (RUCA), DIENST ALGEBRA, GROENENBORGERLAAN 171, B-2020 ANTWERP, BELGIUM