ISRAEL JOURNAL OF MATHEMATICS 139 (2004), 183–188

CHOOSING ROOTS OF POLYNOMIALS SMOOTHLY, II

BY

ANDREAS KRIEGL

Institut für Mathematik, Universität Wien
Strudlhofgasse 4, A-1090 Wien, Austria
e-mail: Andreas.Kriegl@univie.ac.at

AND

MARK LOSIK*

Saratov State University
ul. Astrakhanskaya, 83, 410026 Saratov, Russia
e-mail: losik@info.sgu.ru

AND

PETER W. MICHOR*

Institut für Mathematik, Universität Wien
Strudlhofgasse 4, A-1090 Wien, Austria
and
Erwin Schrödinger Institut für Mathematische Physik
Boltzmanngasse 9, A-1090 Wien, Austria
e-mail: Peter.Michor@esi.ac.at

ABSTRACT

We show that the roots of any smooth curve of polynomials with real roots only can be parametrized twice differentiable (but not better).

In [1] we claimed that there exists a smooth curve of polynomials of degree 3 for which no C^1-parametrization of the roots exists. Unfortunately there was an error in the calculation of b_3 and we have been informed by Jacques Chaumat and Anne-Marie Chollet in June 2001 about that and the related papers [2], [5].

* M. L. and P. W. M. were supported by 'Fonds zur Förderung der wissenschaftlichen Forschung, Projekt P 14195 MAT'.
Received October 2, 2002

We are now going to repair this mistake and improve at the same time the results of [2]. The smoothness assumptions in the following theorem are certainly not the best possible but in fact we are mainly interested in the case of smooth coefficients.

The conclusion of the theorem is the best possible, since even for the characteristic polynomial of a smooth curve of symmetric matrices there needn't be a differentiable parametrization of the roots with locally Hölderian derivative as the first example in [3] shows.

Let P be a curve defined on some subset $T \subseteq \mathbb{R}$ of monic polynomials $P(t)$ of degree $n \geq 1$ with real roots only. A parametrization of some class of the roots of P is a curve $x\colon T \to \mathbb{R}^n$ of that class such that for each $t \in T$ the values $x_1(t), \dots, x_n(t)$ are the roots of $P(t)$ with correct multiplicity.

THEOREM: *Consider a continuous curve of polynomials*

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t), \quad t \in \mathbb{R},$$

with all roots real. Then there is a continuous parametrization $x = (x_1, \dots, x_n)\colon \mathbb{R} \to \mathbb{R}^n$ of the roots of P. Moreover:

(1) [2], Theorem 1 and Theorem 2. *If all coefficients a_i are of class C^n then the parametrization $x\colon \mathbb{R} \to \mathbb{R}^n$ may be chosen differentiable with locally bounded derivative.*

(2) *If all a_i are of class C^{2n} then any differentiable parametrization $x\colon \mathbb{R} \to \mathbb{R}^n$ is actually C^1.*

(3) *If all a_i are of class C^{3n} then the parametrization $x\colon \mathbb{R} \to \mathbb{R}^n$ may be chosen twice differentiable.*

Proof: The parameterization by order $x_1(t) \leq \cdots \leq x_n(t)$ is continuous; see, e.g., [1], 4.1. We prove (2) and (3), and we use the proof of theorem 4.3 in [1]. First we replace x by $x + \frac{1}{n}a_1(t)$, and consequently assume without loss that $a_1 = 0$.

As noted in the proof of 4.3 in [1] the multiplicity lemma [1], 3.7 remains true in the C^m-case for $m \geq n$ in the following sense, with the same proof:

If $a_1 = 0$ then the following two conditions are equivalent:

(1) $a_k(t) = t^k a_{k,k}(t)$ *for a C^{m-k}-function $a_{k,k}$, for all $2 \leq k \leq n$.*

(2) $a_2(t) = t^2 a_{2,2}(t)$ *for a C^{m-2}-function $a_{2,2}$.*

Proof of (2): Let all a_i be C^{2n}.

Then we choose a fixed t, say $t = 0$.

If $a_2(0) = 0$ then it vanishes of second order at 0: if it vanishes only of first order then $\tilde{\Delta}_2(P(t)) = -2na_2(t)$ (see [1], 3.1) would change sign at $t = 0$, contrary to the assumption that all roots of $P(t)$ are real, by [1], 3.2. Thus $a_2(t) = t^2 a_{2,2}(t)$, so by the variant of the multiplicity lemma described above we have $a_k(t) = t^k a_{k,k}(t)$ for C^n-functions $a_{k,k}$, for $2 \leq k \leq n$. We consider the following C^n-curve of polynomials:

$$P^1(t)(z) = z^n + a_{2,2}(t)z^{n-2} - a_{3,3}(t)z^{n-3} \cdots + (-1)^n a_{n,n}(t).$$

Then $P(t)(tz) = t^n\, P^1(t)(z)$ and hence $z \mapsto t\,z = x$ gives for $t \neq 0$ a bijective correspondance between the roots z of $P^1(t)$ and the roots x of $P(t)$ with correct multiplicities. Moreover, parametrizations z which are continuous at $t = 0$ correspond to parametrizations x which are differentiable at $t = 0$. By (1) we may choose the parametrization $z = (z_1, \ldots, z_n)$ differentiable with locally bounded derivative. Then the corresponding parametrization $t \mapsto x(t) := t\,z(t)$ is differentiable with derivative $x'(t) = t\,z'(t) + z(t)$ which is continuous at $t = 0$ with $x'(0) = z(0)$.

If $a_2(0) \neq 0$ then we use the splitting lemma [1], 3.4 for the C^{2n}-case: We may factor $P(t) = P_1(t) \cdots P_k(t)$ for t in a neighborhood of 0 and some $k > 1$ where the P_i have again C^{2n}-coefficients and where each $P_i(0)$ has all roots equal to, say, c_i, and where the c_i are distinct. By the argument above applied to each P_i separately, there is a differentiable parametrization $x = (x_1, \ldots, x_n)$ of roots whose derivative x' is continuous at $t = 0$. Moreover, if $P_i(0)(x_j(0)) = 0$ then $x_j'(0)$ is a root of the polynomial $P_i^1(0)$ which depends only on P_i. We shall use this for arbitrary t below.

Claim: *Any differentiable parametrization $y = (y_1, \ldots, y_n)$ of the roots of P has y' continuous at $t = 0$.* Let $i \in \{1, \ldots, n\}$. For $t_m \to 0$ there are $k_m \in \{1, \ldots, n\}$ such that $y_i(t_m) = x_{k_m}(t_m)$. Choose a subsequence of the t_m again denoted t_m such that $y_i(t_m) = x_k(t_m)$ for some fixed k and all m. By the argument above then we also have $y_i'(t_m) = x_{j_m}'(t_m)$ for some j_m with $x_{j_m}(t_m) = x_k(t_m) = y_i(t_m)$. Passing again to a subsequence we find a fixed j such that $y_i(t_m) = x_j(t_m)$ and $y_i'(t_m) = x_j'(t_m)$. Then

$$\begin{aligned} y_i(0) &= \lim_m y_i(t_m) = \lim_m x_j(t_m) = x_j(0), \\ y_i'(0) &= \lim_m \frac{y_i(t_m) - y_i(0)}{t_m} = \lim_m \frac{x_j(t_m) - x_j(0)}{t_m} = x_j'(0), \end{aligned}$$

and so $y_i'(t_m) = x_j'(t_m) \to x_j'(0) = y_i'(0)$.

Thus any differentiable parametrization of the roots of P (which exists by (1)) is indeed C^1, and (2) is proved.

Proof of (3): Let all a_i be C^{3n}. Remember that $a_1 = 0$.

(a) Choose a fixed t, say $t = 0$. If $a_2(0) = 0$ then we consider again the polynomials $P^1(t)$, which now form a C^{2n}-curve. By (2) its roots can be parametrized by a C^1-curve $t \mapsto z(t) = (z_1(t), \dots, z_n(t))$. The $x(t) = t\,z(t)$ are then again the roots of $P(t)$, now with continuous derivative $x'(t) = t\,z'(t) + z(t)$ which is differentiable at $t = 0$ with $x''(0) = 2\,z'(0)$.

We show by induction on n that for fixed open intervals $I \subseteq \mathbb{R}$ there exists a twice differentiable parametrization y of the roots of P on I.

Let $t_0 \in I$ be such that $a_2(t_0) \neq 0$. By the splitting lemma [1], 3.4 for the C^{3n}-case we may factor $P(t) = P_1(t) \cdots P_k(t)$ for some $k > 1$ and all t in a neighborhood $I_1 \subseteq I$ of t_0 where the $P_i(t)$ have again C^{3n}-coefficients and where each $P_i(t_0)$ has all roots equal to, say, c_i, and where the c_i are distinct. By induction there is on I_1 a twice differentiable parametrization of the roots of each P_i. Note that for $n = 1$ the root equals the (single) coefficient.

Let now $a_2(t) \neq 0$ for all $t \in I$. We consider twice differentiable parametrizations of the roots defined on open subintervals $I_1 \subseteq I$. Obviously we may apply Zorn's lemma to obtain a twice differentiable parametrization on some maximal open subinterval I_1. Suppose for contradiction that $I \supsetneq I_1$ and let the, say, right endpoint t_0 of I_1 belong to I. Then there is a twice differentiable parametrization y on I_1 and, since $a_2(t_0) \neq 0$, a twice differentiable parametrization x in a neighborhood of t_0. Let $t_m \nearrow t_0$. For every m there exists a permutation π of $\{1, \dots, n\}$ such that $y_{\pi(i)}(t_m) = x_i(t_m)$ for all i. By passing to a subsequence, again denoted t_m, we may assume that the permutation does not depend on m. By passing again to a subsequence we may also assume that $y'_{\pi(i)}(t_m) = x'_i(t_m)$ and then again for a subsequence that $y''_{\pi(i)}(t_m) = x''_i(t_m)$ for all i and all m. So we may paste $(y_{\pi(i)}(t))_i$ for $t < t_0$ with $x(t)$ for $t \geq t_0$ to obtain a twice differentiable parametrization on an interval larger than I_1, a contradiction.

Now we consider the closed set

$$E = \{t \in I \colon a_2(t) = 0\} = \{t \in I \colon x_1(t) = \cdots = x_n(t)\}.$$

Then $I \setminus E$ is open, thus a disjoint union of open intervals on which we have a twice differentiable parametrization x of the roots by the previous paragraph.

Consider next the set E' of all accumulation points of E. Then $I \setminus E' = (I \setminus E) \cup (E \setminus E')$ is again open and thus a disjoint union of open intervals, and for each point $t_0 \in E \setminus E'$, i.e. isolated point of E, we have a twice differentiable local parametrization of roots $y_i(t)$ for $t \neq t_0$ (left and right of t_0), and we have a local C^1 parametrization $x_k(t)$ for t near t_0 which is twice differentiable at t_0, by argument (a). Clearly $y_i(t) \to x_1(t_0) = \cdots = x_n(t_0)$ for $t \to t_0$.

For $t_m \searrow t_0$, by passing to a subsequence, we may assume that $y_i'(t_m) = x'_{\pi(i)}(t_m) \to x'_{\pi(i)}(t_0)$. Thus $y_i'(t)$ has at most $x_1'(t_0), \dots, x_n'(t_0)$ as cluster points for $t \searrow t_0$. Since y_i' satisfies the intermediate value theorem, $y_i'(t)$ converges for $t \searrow t_0$, with limit $x'_{\pi(i)}(t_0)$, since it does so along a sequence t_m as above. By renumbering the y_i to the right of t_0 we may assume that $i = \pi(i)$. Similarly for the left side of t_0. Then $y_i'(t) \to x_i'(t_0)$ for $t \to t_0$, so y_i is C^1 near t_0 and still twice differentiable off t_0.

In order to get twice differentiability at t_0 also, we consider again the situation at the beginning of the last paragraph. Then we have

$$\frac{y_i'(t_m) - y_i'(t_0)}{t_m - t_0} = \frac{x'_{\pi(i)}(t_m) - x'_{\pi(i)}(t_0)}{t_m - t_0} \to x''_{\pi(i)}(t_0)$$

so that $(y_i'(t) - y_i'(t_0))/(t - t_0)$ has at most $\{x_j''(t_0) : x_j'(t_0) = y_i'(t_0)\}$ as cluster points for $t \searrow t_0$. Since it satisfies the intermediate value theorem it converges for $t \searrow t_0$, with limit $x''_{\pi(i)}(t_0)$, since it does so along a sequence t_m as just used. Similarly for the left-handed second derivative. Thus we may renumber those y_i for which the $y_i'(t_0)$ agree, to the right of t_0 in such a way that the (one sided) second derivatives agree. Then the (twice) renumbered y_i are twice differentiable also at t_0.

Thus we have a twice differentiable parametrization of roots on the open set $I \setminus E'$.

Now let $t_0 \in E'$, i.e. an accumulation point of E. Let F be the set of all $t \in I$ where $x_1(t) = \cdots = x_n(t)$ and $x_1'(t) = \cdots = x_n'(t)$. Then $t_0 \in F$ since each $x_i'(t_0)$ may be computed using only points in E. Let F' be the set of all accumulation points of F. Thus $E' \subseteq F = (F \setminus F') \cup F' \subseteq E$.

Let first $t_0 \in F \setminus F'$, i.e. an isolated point in F. Then again we have a local twice differentiable parametrization $t \mapsto y(t)$ of the roots for $t \neq t_0$ (left and right of t_0), since near t_0 there are only points in $I \setminus E'$. We still have a local C^1 parametrization x near t_0 which is twice differentiable at t_0, by the argument above. As above we can find a twice differentiable parametrization y of the roots on the open set $(I \setminus E') \cup (F \setminus F')$.

Finally, let $t_0 \in F'$, i.e. an accumulation point in F. We use again parametrizations x near t_0, and y as above. Then all $x_i(t_0)$ agree, all $x_i'(t_0)$ agree, and even all $x_i''(t_0)$ agree. We extend each y_i from $(I \setminus E') \cup (F \setminus F')$ by these single functions on F' to the whole of $(I \setminus E') \cup (F \setminus F') \cup F' = (I \setminus E') \cup F = I$. We have to check that then each y_i is twice differentiable at t_0. For $t_m \to t_0$ we have, by passing to subsequences,

$$y_i(t_m) = x_j(t_m) \to x_j(t_0) = x_i(t_0) = y_i(t_0),$$
$$\frac{y_i(t_m) - y_i(t_0)}{t_m - t_0} = \frac{x_j(t_m) - x_j(t_0)}{t_m - t_0} \to x_j'(t_0) = x_i'(t_0),$$
$$\frac{y_i'(t_m) - y_i'(t_0)}{t_m - t_0} = \frac{x_j'(t_m) - x_j'(t_0)}{t_m - t_0} \to x_j''(t_0) = x_i''(t_0). \quad \blacksquare$$

References

[1] D. Alekseevsky, A. Kriegl, M. Losik and P. W. Michor, *Choosing roots of polynomials smoothly*, Israel Journal of Mathematics **105** (1998), 203–233, arXiv: math.CA/9801026.

[2] M. D. Bronshtein, *Smoothness of polynomials depending on parameters*, Sibirskii Matematicheskii Zhurnal **20** (1979), 493–501 (Russian); English transl.: Siberian Mathematical Journal **20** (1980), 347–352.

[3] A. Kriegl and P. W. Michor, *Differentiable perturbation of unbounded operators*, Mathematische Annalen, to appear, arXiv: math.FA/0204060.

[4] Y. Ohya and S. Tarama, *Le problème de Cauchy à caracteristiques multiples dans la classe de Gevrey (coefficients Hölderiens en t)*, in *Hyperbolic Equations and Related Topics*, Proceedings of the Taniguchi International Symposium, Katata and Kyoto/Jap. 1984, Academic Press, Boston, 1986, pp. 273–306.

[5] S. Wakabayashi, *Remarks on hyperbolic polynomials*, Tsukuba Journal of Mathematics **10** (1986), 17–28.