The path to the zeros of a polynomial

by Bernard Beauzamy

Abstract. – Let $P(z)$ be a polynomial in one complex variable, with complex coefficients, and let $z_1, \dots, z_n$ be its zeros. Assume, by normalization, that $P(0) = 1$. The *direct path* from 0 to the root z_j is the set $\{P(tz_j),\ 0 \le t \le 1\}$. We are interested in the *altitude* of this path, which is $|P(tz_j)|$. We show that there is always a zero towards which the direct path declines near 0, which means $|P(tz_j)| < |P(0)|$ if t is small enough. However, starting with degree 5, there are polynomials for which no direct path constantly remains below the altitude 1.

April 1996

In application of contract E.T.C.A./C.R.E.A. 95.01.181, Ministry of Defense, France

Although many algorithms exist which, starting at a given point z_0 in the complex plane, find the zeros of a given polynomial, none of them is completely satisfactory, neither in theory nor in practice.

In practice, the most widely used are variations upon Newton's method, Traub-Jenkins algorithm [3] and Schönhage's algorithm [5]. They work satisfactorily in a number of cases, but may fail if the degreee of the polynomial is too high or if the roots are too clustered. Another algorithm, due to the author [2], has not yet been practically implemented, but suffers certainly from these defects and/or some others.

This is due to the fact that the theory behind all these algorithms is not well-understood. In what cases (that is : for what polynomials ?) is one algorithm better than the other ? What situations (i.e. dispositions of the zeros) will slow down any of them ? At present, such questions are quite unclear.

Indeed, there is a strong lack of general information regarding the zeros of a polynomial. By "general information", I mean quantitative data, depending only on the coefficients, and satisfied by any polynomial. As an example, let me mention an estimate for the radius of the smallest disk, centered at 0, and containing at least a zero. This radius r satisfies

$$r \ \le\ \sqrt{n[P]^2 - 1}, \tag{1}$$

where $P(z) = \sum_0^n a_j z^j$ and $[P] = (\sum_{j=0}^n |a_j|^2 / \binom{n}{j})^{1/2}$ is Bombieri's norm of P (see [1]). Many other estimates of r exist in the literature, for instance due to Cauchy (see Marden [4]), sometimes weaker than (1), sometimes stronger, but nobody knows for which polynomials which one is best.

Another general information was given by Smale [6]. Let $P(z)$ be a polynomial, $z_1, \dots, z_n$ its roots, all different from 0. Then, for one of the zeros, say z_j, one has

$$\frac{1}{|z_j|}\ \left| \int_0^{z_j} P(\zeta) d\zeta \right| \ \le\ 4|P(0)|, \tag{2}$$

and S. Smale asks if 4 cannot be replaced by a smaller number.

This question can be interpreted in a more general setting, meeting exactly what I described at the beginning of this introduction : if you start at a given point, say $z = 0$, and if you look at $|P(z)|$ when z moves from 0 to one of the zeros, z_j, what can you say about the path ? For instance, what is its length and what is its maximal altitude ? In this short note, we have concentrated ourselves on the second question (the altitude), which looks easier than the first.

Let now $P(z) = (z - z_1) \cdots (z - z_n)$ be a polynomial, normalized with leading coefficient 1. We call "direct path to the zero z_j" the set $\{P(tz_j),\ 0 \le t \le 1\}$. We start at 0, and we will investigate the altitude of the path, which is $\max_{0 \le t \le 1} \ |P(tz_j)|$.

Theorem. *– 1. For every polynomial P with $P(0) \neq 0$, there is a root z_j for which the direct path towards this root is initially declining, which means that there is an $\epsilon > 0$ (depending on P and on the zero), for which*

$$|P(tz_j)| < |P(0)|, \tag{3}$$

for all t, $0 < t < \epsilon$.

– 2. However, there exist polynomials (even with all roots on the unit circle) for which no direct path stays under the horizontal plane of altitude $|P(0)|$, which means that, for such P's :

$$\max_{0 \le t \le 1} |P(tz_j)| > |P(0)|, \tag{4}$$

for every $j = 1, \dots, n$.

Proof of the theorem.

The first part will follow from Taylor's formula. Since $P(z)$ is an analytic function, it is clear that, starting at any point, there are directions (totalizing an angle of π) at which $|P(z)|$ diminishes, but we have to prove that one of these directions is the direction of a root.

– Let us first consider the case where $P'(0) \neq 0$. We write $P(z) = \sum_0^n a_j z^j$, with $a_0 \neq 0,\ a_1 \neq 0$. For any zero z_j, we have

$$|P(tz_j)|^2 \sim |a_0|^2 + 2t \ Re(a_0 \overline{a}_1 \overline{z}_j)$$

when $t \to 0$. So all we have to show is that there is a zero z_j for which

$$Re \ (a_0 \overline{a}_1 \overline{z}_j) < 0. \tag{5}$$

This is equivalent to :

$$Re \ \left((-1)^n \ (\prod_{i=1}^n z_i) \ (-1)^{n-1} \ (\prod_{i=1}^n \overline{z}_i) \ (\sum_{i=1}^n \frac{1}{\overline{z}_i}) \overline{z}_j \right) \ < \ 0,$$

or

$$Re \ (z_j (\sum_{i=1}^n \frac{1}{z_i}) \ > \ 0. \tag{6}$$

But (6) is clear : by a proper rotation of the x axis (which changes nothing to the problem), we may assume $\sum_i 1/z_i$ to be real positive. Then, one of the $Re(1/z_j)$ has to be real positive, and so is $Re(z_j)$, which proves (6).

– Let us now look at the general case : $a_0 \neq 0,\ a_1 = \cdots = a_{k-1} = 0,\ a_k \neq 0 \ (k \le n)$, which is more difficult.

Then, for $z \neq 0$, t small enough,

$$|P(tz)|^2 \sim |a_0|^2 + 2t^k \ Re(a_0 \overline{a}_k \overline{z}^k), \tag{7}$$

and we want to show that there is a zero z_j for which

$$Re(a_0 \overline{a}_k \overline{z}^k) < 0. \tag{8}$$

This condition can be rewritten, with $\beta_j = 1/z_j$:

$$Re\left((-1)^n \; (\prod_{i=1}^{n} z_i) \; (-1)^{n-k} \; (\prod_{i=1}^{n} \overline{z}_i) \sum_{i_1<\cdots<i_k} \overline{\beta}_{i_1} \cdots \overline{\beta}_{i_k} \overline{z}_j^k \right) < 0,$$

or

$$(-1)^k \; Re \left(z_j^k \sum_{i_1<\cdots<i_k} \beta_{i_1} \cdots \beta_{i_k} \right) < 0. \tag{9}$$

We let $S_1 = \sum_{i=1}^n \beta_i$, $S_2 = \sum_{i_1<i_2} \beta_{i_1}\beta_{i_2}$, and so on until $S_k = \sum_{i_1<\cdots<i_k} \beta_{i_1} \cdots \beta_{i_k}$.

The quantity S_1^k can be written :

$$S_1^k = \sum_{i=1}^{n} \beta_i^k + R(S_1, \ldots, S_{k-1}) + C \; S_k, \tag{10}$$

where R is a polynomial in $S_1, \ldots, S_{k-1}$ with no constant term, and C is a constant (independent of P) which we now determine.

To this aim, we take $\beta_j = e^{2ij\pi/k}$, $j = 0, \ldots, k-1$. Then $S_1 = \cdots = S_{k-1} = 0$, and

$$S_k = \beta_1 \cdots \beta_k = e^{2i\pi(1+2+\cdots+k)/k} = (-1)^{k+1},$$

so $C = (-1)^k/k$, and formula (10) becomes

$$S_1^k = \sum_{i=1}^{n} \beta_i^k + R(S_1, \ldots, S_{k-1}) + \frac{(-1)^k}{k} S_k. \tag{11}$$

Since we assumed $a_1 = \cdots = a_{k-1} = 0$, we get $S_1 = \cdots = S_{k-1} = 0$, and therefore

$$\sum_{i=1}^{n} \beta_i^k \; = \; \frac{(-1)^{k+1}}{k} \; S_k. \tag{12}$$

In order to prove (9), we may assume S_k to be real positive (by a global rotation of the picture, as we did previously), and so $\sum_{i=1}^n (-1)^{k+1} \beta_i^k$ is real positive.

So, there must be an index j for which $Re((-1)^{k+1}\beta_j^k) > 0$.

But this is equivalent to

$$Re \left(\frac{(-1)^{k+1}}{z_j^k} \right) \; > \; 0, \tag{13}$$

which proves (9) and finishes the proof of the first part of the Theorem.

The proof of the second part will rely upon the ideas we just presented. Let us come back to the case $a_1 \neq 0$. We have seen in formula (5) that the direct path towards the root z_j is locally declining near the origin if $Re(z_j \sum_{i=1}^n 1/z_i) > 0$.

An obvious remark is that this estimate may very well hold for one zero only : indeed, we may have $\sum 1/z_i$ real positive, but only one of the z_i's has positive real part. This is the case, for instance, for a polynomial of the form $(z-e^{i\theta})^k (z-e^{-i\theta})^k (z-1)$, if $\theta > \pi/2$ is close enough to $\pi/2$. For such a polynomial, the direct paths towards $e^{i\theta}$ and $e^{-i\theta}$ will be locally climbing near 0. But the value of $|P|$ on the path to 1, say $|P(1/2)|$, satisfies $|P(1/2)| \geq \frac{1}{2}(\frac{5}{4})^k$, and this is larger than 1 if k is large enough. Precisely, we have :

Proposition 2. – *Let* $P(z) = (z-e^{i\theta})^4(z-e^{-i\theta})^4(z-1)$, *where* $-1/8 < \cos\theta < 0$. *Then the direct paths towards* $e^{i\theta}$ *and* $e^{-i\theta}$ *are locally climbing near* 0, *and* $|P(1/2)| > \frac{1}{2}(\frac{5}{4})^4 > 1.22$.

Proof of Proposition 2. – We have :

$$|P(te^{i\theta})| \ = \ (1-t)^4 \ |1-te^{2i\theta}|^4 \ |1-te^{i\theta}|$$

and so

$$|P(te^{i\theta})|^2 \ = \ (1-t)^8 \ (1-2t\cos 2\theta + t^2)^4 \ (1-2t\cos\theta + t^2) \ \sim \ 1-2t(4+4\cos 2\theta + \cos\theta),$$

and $4+4\cos 2\theta + \cos\theta = (8\cos\theta+1)\cos\theta$; the result follows. The largest slope at the origin is obtained with $\cos\theta = -1/16$; it gives $|P(te^{i\theta})|^2 \sim 1+t/16$.

In fact, this phenomenon is quite general, and appears at any degree, starting at $n=5$. Indeed, the polynomial

$$P(z) \ = \ (z-e^{i\theta})^2(z-e^{-i\theta})^2(z-1),$$

with $\cos\theta = -1/6$ satisfies :

$$|P(te^{i\theta})|^2 \ = \ |P(te^{-i\theta})|^2 \ = \ 1.001168.....$$

for $t = 0.0209.....$, and

$$|P(t)|^2 \ = \ 1.0122....$$

for $t = 0.4527.....$

On the other hand, it can be shown that for any polynomial of degree 4 or less, there is always a direct path which stays belows the plane of altitude $|P(0)|$ (and so, for such a polynomial, the 4 in Smale's formula can be reduced to 1).

Finally, we observe that there is no absolute constant C such that, for any polynomial P,

$$\min_{z, P(z)=0} \ \max_{0\le t\le 1} \ |P(tz)| \ \le \ C|P(0)|. \qquad (14).$$

Indeed, if P_5 is the polynomial of degree 5 we just computed, we have $P_5(0)=1$, and

$$\min_{z, P_5(z)=0} \ \max_{0<t<1} \ |P_5(tz)| \ \le \ \alpha > 1,$$

so

$$\min_{z, P_5(z)=0} \ \max_{0<t<1} \ |P_5^k(tz)| \ = \ \alpha^k \ \to +\infty, \ \text{ when } \ k \to +\infty,$$

which shows that there cannot be in (14) a bound C independent of the degree ; moreover, the bound must be at least exponential in the degree.

References

[1] Beauzamy, B. : Products of Polynomials and a priori estimates for coefficients in polynomial decompositions : a sharp result. *Journal of Symbolic Computation*, vol. 13, 1992, pp. 463–472.

[2] Beauzamy, B. : Finding the roots of polynomial equations : an algorithm with linear command. *To appear.*

[3] Jenkins, M.A. – Traub, J.F. : Principles for testing polynomial zero finding programs. *ACM Transaction on Mathematical Software*, vol. 1, March 1975, no 1, pp. 2634.

[4] Marden, M. : Geometry of Polynomials. *A.M.S. Math. Surveys*, 3rd edition, 1985.

[5] Schönhage, A. : Equation solving interms of computational complexity. *Proceedings International Congress of Mathematics*, Berkeley, 1987.

[6] Smale, S : The fundamental theorem of algebra and complexity theory. *Bull. A.M.S.*, vol 4, 1981, no 1, pp. 1–36.