Critical points of polynomials*

Vilmos Totik

November 11, 2020

Abstract

This paper is devoted to the problem of where the critical points of a polynomial are relative to their zeros. Classical and new developments are surveyed along with illustrative examples. The paper finishes with a short proof of the sector theorem of Sendov and Sendov.

This paper surveys some of the results regarding the location of the critical points of polynomials. A short proof will also be given for the beautiful recent sector theorem of B. Sendov and H. Sendov.

The topic is very old, the most classical references are [30] by E. B. Van Vleck, [14] by M. Marden, [17] by Q. I. Rahman and G. Schmeisser and [24] by T. Sheil-Small. We shall however, also touch many newer developments that are not included in those works. We also mention the recent survey paper [21] by T. Richards that discusses some of the topics to be dealt with below.

1 The Gauss-Lucas theorem

Let us start with the classical formulation.

Theorem 1 (Gauss, Lucas) *If P is a non-constant polynomial, then the convex hull of its zeros contains the critical points of P, i.e. the zeros of P'.*

This is an easy consequence of an observation by C. F. Gauss (cf. [7]) from around 1836 describing the critical points as the equilibrium points in a field generated by unit charges placed at the zeros counting multiplicity (see below), and it was was explicitly stated and proved by F. Lucas [12] in 1874.

The Gauss-Lucas theorem is often stated in the form that if K is a (closed) convex set that contains all zeros of P, then K contains all zeros of P'.

The proof is simple: since K is the intersection of half-planes, it is sufficient to show the claim when K is a half-plane, which we may assume to be $K =$

*AMS Classification: 26C10, 31A15; Key words: zeros, critical points, polynomials

$\{z \mid \Re z \leq 0\}$. Let $z_1, \ldots, z_n$ be the zeros of P. Thus, $\Re z_j \leq 0$, hence if $z \notin K$, i.e. $\Re z > 0$, then $\Re(z - z_j) > 0$ for all j, and so

$$\Re \frac{1}{z - z_j} = \frac{\Re(z - z_j)}{|z - z_j|^2} > 0.$$

But then, since

$$\frac{P'(z)}{P(z)} = \sum_{j=1}^{n} \frac{1}{z - z_j}, \tag{1}$$

we have

$$\Re \frac{P'(z)}{P(z)} = \sum_{j=1}^{n} \Re \frac{1}{z - z_j} > 0,$$

showing that $P'(z) \neq 0$.

■

The proof easily gives that if the zeros of P are not collinear, then a critical point lies in the interior of the convex hull of the zeros unless it is a zero of P of multiplicity ≥ 2.

Furthermore, the same proof shows that if $r_1, \ldots, r_n$ are non-negative numbers not all zero, then all zeros of

$$Q(z) = \sum_{j=1}^{n} r_j \prod_{k \neq j} (z - z_k)$$

lie in the convex hull of $z_1, \ldots, z_n$ (c.f. [30, p. 648], [3]). The original Gauss-Lucas statement is the $r_1 = \cdots = r_n = 1$ special case.

Gauss' formula (1) gives rise to an electrostatic interpretation of the critical points. Indeed, place a unit positive charge to every zero counting multiplicity (so at a zero we place charge m if m is its multiplicity), and consider the potential field of these charges provided the attractive/repelling force is proportional with the reciprocal of the distance (on the plane this is the version of Coulomb's law — in the plane the potential field is generated by the logarithmic kernel). Note that for a unit positive charge placed at a point z the charge placed at the zero z_j exercises the force

$$c \overline{\frac{1}{z - z_j}},$$

where $\bar{\cdot}$ denotes complex conjugation and c is a fixed, universal constant (Coulomb constant). Hence, formula (1) shows that z is a critical point of P precisely if it is an equilibrium point in that field, i.e. if the total force at z is zero:

$$c \sum_{j=1}^{n} \overline{\frac{1}{z - z_j}} = 0.$$

We close this section by giving another proof of the Gauss-Lucas theorem due to T. Richards (see the proof of [19, Theorem 2.1]). Suppose to the contrary that there is a w not in the convex hull of the zeros which is a critical point of P. Then w and the convex hull can be separated by a line ℓ. We may assume that ℓ is the imaginary axis, all zeros z_j lie in $\{z \mid \Re z < 0\}$ and $\Re w > 0$. If the multiplicity of w in P' is $k \geq 1$, then in a neighborhood of w the level set $S := \{z \mid |P(z)| = |P(w)|\}$ consists of $k+1$ analytic arcs such that their tangent lines at w divide the the plane into $(2k+2)$ sectors of vertex angle $2\pi/(2k+2)$. By rotating the zeros of P about w a little we may assume that neither of the just mentioned tangent lines is horizontal. Then every horizontal half-line $\{iy + t \,|\, t \geq 0\}$ that is close to w intersects S in at least two different points. But that is impossible, since $|P(z)|$, $z = iy + t$, is increasing along every such horizontal line (each $|z - z_j|$ with $\Re z_j < 0$ increases as $t \geq 0$ increases in $z = iy + t$).

■

2 Higher derivatives

Let $Z(P) = \{z_1, \ldots, z_n\}$ denote the zero set of $P(z) = a_n z^n + \cdots + a_0$ $(a_n \neq 0)$ and $\mathcal{C}(Z(P))$ its convex hull. This is a (possibly degenerate) polygon with center of mass

$$\frac{z_1 + \cdots + z_n}{n} = -\frac{a_{n-1}}{n a_n}.$$

By the Gauss-Lucas theorem $\mathcal{C}(Z(P')) \subset \mathcal{C}(Z(P))$, and here the center of mass of $\mathcal{C}(Z(P'))$ is again $-a_{n-1}/na_n$ because

$$P'(z) = n a_n z^{n-1} + (n-1) a_{n-1} z^{n-2} + \cdots.$$

Now applying this to P' and then to P'' etc., we obtain that

$$\mathcal{C}(Z(P)) \supset \mathcal{C}(Z(P')) \supset \mathcal{C}(Z(P'')) \supset \cdots \supset \mathcal{C}(Z(P^{(n-1)})) = \left\{ -\frac{a_{n-1}}{n a_n} \right\} \qquad (2)$$

are shrinking polygons with the same center of mass.

This has the following converse, see [8] by T. Genchev and B. Sendov.

Theorem 2 (Genchev-Sendov) *Let $L : \mathbf{C}(z) \to \mathbf{C}(z)$ be a linear operator from $\mathbf{C}(z)$ (the set of continuous functions on $\mathbf{C}$) into itself such that if P is a non-constant polynomial and $L(P) \neq 0$, then $\mathcal{C}(P)$ contains $\mathcal{C}(L(P))$. Then either L is a linear functional, or there is a $c \neq 0$ and a k such that $L(P) = cP^{(k)}$ for all polynomials P.*

For a related result when $\mathcal{C}(L(P)) \subseteq \mathcal{C}(P)$ is replaced by the assumption that the diameter $\operatorname{diam}\mathcal{C}(L(P))$ of $\mathcal{C}(L(P))$ is at most as large as $\operatorname{diam}\mathcal{C}(P)$ see the paper [15] by N. Nikolov and B. Sendov.

It is a remarkable fact that in the decreasing sequence (2) the terms become small for high order derivatives. Indeed, the results of [18] by M. Ravichandran easily imply

Theorem 3 (Ravichandran) *If the degree of P is n, then for $c \geq 1/2$ we have*

$$\operatorname{diam} Z(P^{(cn)}) \leq 2\sqrt{c(1-c)} \cdot \operatorname{diam} Z(P)). \tag{3}$$

In particular, if c is close to 1, then the diameter of $Z(P^{(cn)})$ is much smaller than the diameter of $Z(P)$, and the rate of decrease is universal.

Proof. Let P be of degree n with zeros $z_1, \ldots, z_n$, and let $R(P)$ denote the monic polynomial with zeros $\Re z_1, \ldots, \Re z_n$. Furthermore, let us denote by $\lambda_{\max}(Q)$, $\lambda_{\min}(Q)$ the largest resp. smallest zero of a polynomial Q with real zeros. Corollary 5.4 in [18] states that for any $k \geq 1$

$$\lambda_{\max}(R(P^{(k)})) \leq \lambda_{\max}((R(P))^{(k)}), \qquad \lambda_{\min}(R(P^{(k)})) \geq \lambda_{\min}((R(P))^{(k)}). \tag{4}$$

Furthermore, Lemma 6.1 from [18] claims that if Q is of degree n and has only real zeros, then

$$\lambda_{\max}(Q^{(cn)}) - \lambda_{\min}(Q^{(cn)}) \leq 2\sqrt{c(1-c)} \cdot \Big(\lambda_{\max}(Q) - \lambda_{\min}(Q)\Big).$$

In particular,

$$\lambda_{\max}((R(P))^{(cn)}) - \lambda_{\min}((R(P))^{(cn)}) \leq 2\sqrt{c(1-c)} \cdot \Big(\lambda_{\max}(R(P)) - \lambda_{\min}(R(P))\Big).$$

Putting these together we obtain

$$\lambda_{\max}(R(P^{(cn)})) - \lambda_{\min}(R(P^{(cn)})) \leq 2\sqrt{c(1-c)} \cdot \Big(\lambda_{\max}(R(P)) - \lambda_{\min}(R(P))\Big).$$

What we have obtained is that the length of the vertical projection of $\mathcal{C}(P^{(cn)})$ onto the real line is at most $2\sqrt{c(1-c)}$ times the length of the same projection of $\mathcal{C}(P)$. Of course, then the same is true about the perpendicular projection of these convex hulls onto any line.

Let now d be the diameter of $Z(P^{(cn)})$, and suppose that $w_1, w_2 \in Z(P^{(cn)})$ are two points for which $|w_1 - w_2| = d$. We may assume without loss of generality, that $w_1, w_2 \in \mathbf{R}$ and $w_1 < w_2$. Then the zeros of $P^{(cn)}$ all lie in the vertical strip determined by the lines $x = w_1$ and $x = w_2$. By what we have just shown, the smallest vertical strip that contains all zeros of P must have width $\geq d/2\sqrt{c(1-c)}$, so there are zeros z_1, z_2 of P for which

$\Re z_2 - \Re z_1 \geq d/2\sqrt{c(1-c)}$. But then $\operatorname{diam} Z(P) \geq d/2\sqrt{c(1-c)})$, and this is what we wanted to prove.

■

The paper [18] also contains the example $P(z) = (z^2-1)^m$, $n = 2m \to \infty$, showing that the bound given in Proposition 3 is of the correct order for c lying close to 1. In fact, in this case Vieta's formulae for

$$P^{(cn)}(z) = n(n-1)\cdots(n-cn+1)z^{(1-c)n} - \frac{n}{2}(n-2)(n-3)\cdots(n-cn-1)z^{(1-c)n-2} + \cdots$$

give that

$$\begin{aligned}\sum_{\lambda\in Z(P^{(cn)})} \lambda^2 &= \Big(\sum_{\lambda\in Z(P^{(cn)})} \lambda\Big)^2 - 2\sum_{\lambda,\theta\in Z(P^{(cn)}),\ \lambda\neq\theta} \lambda\theta \\ &= 0 + \frac{(n-cn)(n-cn-1)}{n-1} = n(1-c)^2 + O(1),\end{aligned}$$

and since the left-hand side is the sum of $n(1-c)$ numbers, it follows that $\lambda^2 \geq 1-c+O(1/n)$ for the largest λ^2. Now we can deduce from the symmetry of the zeros that

$$\begin{aligned}\operatorname{diam} Z(P^{(cn)}) &= \lambda_{\max}(P^{(cn)}) - \lambda_{\min}(P^{(cn)}) \geq 2\sqrt{1-c} + O(1/n) \\ &= \sqrt{1-c}\cdot \operatorname{diam} Z(P) + O(1/n),\end{aligned}$$

showing that the decrease of the diameter of the zero set after cn derivation is not smaller than $\sqrt{1-c} + O(1/n)$.

The same example explains why c in Theorem 3 must be at least 1/2: for $k < n/2$ the diameter of $Z(P^{(k)})$ is the same as the diameter of $Z(P)$.

3 The theorem of Malamud and Pereira

Let P be a polynomial of degree n, let $z_1, \ldots, z_n$ be its zeros and $\xi_1, \ldots, \xi_{n-1}$ the zeros of P_n'. The Gauss-Lucas theorem states that each ξ_j is a convex linear combination of $z_1, \ldots, z_n$. A remarkable extension was proved independently in [13] and [16] by S. M. Malamud and R. Pereira. To state their result recall that an $(n-1)\times n$ size $\mathcal{A} = (a_{ij})$ matrix is doubly stochastic if

- $a_{ij} \geq 0$,
- each row-sum equals 1, and
- each column-sum equals $(n-1)/n$.

Let

$$\mathbf{Z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \qquad \Xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_{n-1} \end{pmatrix}$$

With these notations the celebrated Malamud-Pereira theorem can be stated as

Theorem 4 (Malamud, Pereira) *There is a doubly stochastic matrix $\mathcal{A}$ such that $\Xi = \mathcal{A}\mathbf{Z}$.*

This has a very strong immediate consequence:

Corollary 5 *If $\varphi : \mathbf{C} \to \mathbf{R}_+$ is convex, then*

$$\frac{1}{n-1}\sum_{j=1}^{n-1}\varphi(\xi_j) \le \frac{1}{n}\sum_{k=1}^{n}\varphi(z_k). \tag{5}$$

Convexity is meant in the classical sense that

$$\varphi(\alpha z + (1-\alpha)w) \le \alpha\varphi(z) + (1-\alpha)\varphi(w)$$

for all z, w and $0 < \alpha < 1$. As usual, this implies

$$\varphi(a_1 z_1 + \cdots + a_n z_n) \le a_1\varphi(z_1) + \cdots + a_n\varphi(z_n)$$

for $a_j \ge 1$, $\sum_j a_j = 1$. From here the proof of (5) follows in two lines:

$$\frac{1}{n-1}\sum_{j=1}^{n-1}\varphi(\xi_j) \le \frac{1}{n-1}\sum_{j=1}^{n-1}\sum_{k=1}^{n} a_{jk}\varphi(z_k)$$

$$= \frac{1}{n-1}\sum_{k=1}^{n}\varphi(z_k)\sum_{j=1}^{n-1} a_{jk} = \frac{1}{n}\sum_{k=1}^{n}\varphi(z_k).$$

We list a few special cases of (5).

1) For $m \ge 1$

$$\frac{1}{n-1}\sum_{j=1}^{n-1}|\Re\xi_j|^m \le \frac{1}{n}\sum_{k=1}^{n}|\Re z_k|^m, \qquad m \ge 1.$$

For $m = 1$ this was proved by P. Erdős and I. Niven [6] and simultaneously by N. G. de Bruijn and T. A. Springer [1].

2) If $m \geq 1$, then

$$\frac{1}{n-1}\sum_{j=1}^{n-1}|\xi_j|^m \leq \frac{1}{n}\sum_{k=1}^{n}|z_k|^m.$$

For integer m this is due to de Bruijn and Springer [1]. That same paper conjectured (5), which conjecture was open for about 50 years until the papers of Malamud and Pereira.

3) If all zeros lie in the upper-half plane, then

$$\left(\prod_{k=1}^{n}\Im z_k\right)^{1/n} \leq \left(\prod_{j=1}^{n-1}\Im \xi_j\right)^{1/(n-1)}.$$

4 Asymptotic Gauss-Lucas Theorem

In the Gauss-Lucas theorem there are two essential assumptions:

1. K is convex,
2. all zeros of P lie in K.

Convexity cannot be dropped: if K is closed and not convex, then there are $w_1, w_2 \in K$ such that the line segment connecting w_1 and w_2 lies entirely outside K (except for its endpoints), and then for the polynomial $P(z) = (z - w_1)(z - w_2)$, which has zeros in K, its only critical point (which is the midpoint of segment $\overline{w_1w_2}$) lies outside K. Note, however, that for the case $K = [-2, -1] \cup [1, 2]$ all but (possibly) one critical points of a polynomial lie in K if all of its zeros are from K.

Regarding condition 2. the situation is more dramatic: if one zero of P is allowed to be outside K, then it can happen that all the critical points lie outside K. Indeed, if $K = [-1, 1]$ and $P_0(z) = (z-i)\prod_{j=1}^{n-1}(z-x_j)$ where $x_1, \ldots, x_{n-1} \in [-1, 1]$ are different, then all critical points have positive imaginary part, and hence lie outside K. Simple linear transformation of this example gives a similar example for all convex K. Indeed, pick a point S on the boundary ∂K of K where this boundary (as a plane curve) is differentiable and hence there is a unique tangent line there. We may assume that $S = 0$, the real line is the tangent line to K and K lies in the lower half-plane $\{z \mid \Im z \leq 0\}$. Take a linear transformation $Tz = \varepsilon z + \eta$ with small $\varepsilon > 0$ and η such that $T[-1, 1]$ becomes a horizontal chord I of ∂K lying close to 0, and consider the polynomial $\tilde{P}_0(z) = P_0(T^{-1}(z))$. Then all but (possibly) one of the zeros of $\tilde{P}$ lie in K, and it is easy to see that for small ε (depending on n and the points $x_1, \ldots, x_{n-1} \in [-1, 1]$) all of its $(n - 1)$ critical points lie outside K (just note that the portion of ∂K that lies above the chord I is of distance $o(\varepsilon)$ from I as $\varepsilon \to 0$, while the images of the critical points of P under T lie of distance $\geq c\varepsilon$ above I).

The just given example seems to rule out an analogue of the Gauss-Lucas theorem when some of the zeros of P may lie outside the convex set K. However, while in this case all critical points may lie outside K, most of them must be close to K (an observation of B. Shapiro), and this gives rise to an asymptotic version of the Gauss-Lucas theorem.

In what follows we shall consider polynomials P_n, $n = 1, 2, \ldots$, where the degree of P_n is n. We say that most of the zeros of P_n lie in K if

$$\lim_{n\to\infty} \frac{1}{n}\#\{\text{zeros of } P_n \text{ lying in } K\} = 1,$$

that is if only $o(n)$ of the zeros of P_n lie outside K. With this notation it is true (see [27]) that if K is convex and most of the zeros of P_n, $n = 1, 2, \ldots$, lie in K, then most of the zeros of P_n' lie in any fixed neighborhood of K.

However, in this asymptotic version the convexity of K is not essential, the same statement holds for so-called polynomially convex sets K. To formulate it we need the following definition. Let $K \subset \mathbf{C}$ be a non-empty compact set, and let Ω be the unbounded connected component of $\mathbf{C} \setminus K$. The set $\mathrm{Pc}(K) = \mathbf{C} \setminus \Omega$ is called the polynomially convex hull of K. It is the union of K with all the bounded components of $\mathbf{C} \setminus K$. Now with it we can formulate the following asymptotic Gauss-Lucas theorem (see [28, Corollary 1.9]).

Theorem 6 *If K contains most of the zeros of P_n, $n = 1, 2, \ldots$, then any neighborhood of $\mathrm{Pc}(K)$ contains most of the zeros of P_n'.*

Recall, that P_n' may not have a single zero in $\mathrm{Pc}(K)$, so we must to consider neighborhoods (which may, however, be as small as we wish).

For convex K T. J. Richards conjectured (see [20], [22]) the following quantitative version, in which K_ε denotes the ε-neighborhood of K: *If K is convex, then for $\varepsilon > 0$ there is an $\eta > 0$ (depending on K and ε) such that if a polynomial P_n of degree n has $k \geq (1-\eta)n$ zeros in K, then P_n' has at least $k-1$ zeros in K_ε.*

This conjecture is true, a proof will be published in [29]. This latter paper also gives the bounds $C_1\varepsilon^2 \leq \eta \leq C_2\varepsilon$ for the best constant $\eta = \eta_\varepsilon$. A weaker version (namely when P_n has $k \geq n(1 - c_\varepsilon/\log n)$ zeros in K), was proven in [20] by T. J. Richards and S. Steinerberger.

In the just mentioned quantitative result the number k is close to n. There is no version of the sort that "if K contains a certain portion αn of the zeros, then K_ε will contain some portion $\alpha' n$ of the critical points". Indeed, this completely fails for $\alpha < 1/2$ as is seen from

Example 7 Let $\alpha < 1/2$. If $P_n(z) = z^n - 1$, and K is the square of side-length 2 and with center at the point $1 + 2\varepsilon$, then for small $\varepsilon > 0$ and large n the set K contains at least αn of the zeros of P_n (n-th roots of unity). However, all the critical points are at the origin, so K_ε does not contain a single critical point.

5 Distributions of zeros and critical points

Let us consider again a sequence of polynomials P_n, $n = 1, 2, \ldots$, where P_n is of degree n. If $z_{1n}, \ldots, z_{nn}$ are the zeros of P_n, then the zero counting measure ν_{P_n} is defined as

$$\nu_{P_n} = \frac{1}{n} \sum_{j=1}^{n} \delta_{z_{jn}},$$

where δ_z denotes the Dirac measure (unit mass) at the point z. When we talk about zero distribution, we talk about weak* convergence of the sequence $\{\nu_{P_n}\}$ (or of a subsequence of it): we say that ν_{P_n} tends to the measure μ if

$$\int g d\nu_{P_n} = \frac{1}{n} \sum_{j=1}^{n} g(z_{jn}) \to \int g d\mu, \qquad n \to \infty,$$

for all continuous functions g on $\mathbf{C}$ of compact support. This is equivalent to the fact that for a dense set of disks D (more precisely for all disks D for which $\mu(\partial D) = 0$)

$$\frac{1}{n} \#\{1 \leq j \leq n \,|\, z_{jn} \in D\} \to \mu(D), \qquad n \to \infty.$$

In what follows we shall always assume that μ is of compact support S and of total mass 1 (i.e. only $o(n)$ of the zeros can go to infinity). On applying Theorem 6 to a large disk containing the support S of μ we can see that then only $o(n)$ of the zeros of P_n' can go to ∞, i.e., by Helly's selection theorem, from any subsequence of $\{\nu_{P_n'}\}$ we can select a convergent subsequence, the limit of which we denote by ν. The basic question we are discussing in this section how ν and μ are related.

First we consider some illustrative examples.

Example 8 Suppose that all zeros of P_n lie in $[-1, 1]$. By Rolle's theorem in between any two zeros of P_n there is a zero of P_n' and the multiplicity of a zero is decreased by 1 under differentiation, which easily imply that $\nu = \mu$. In other words, in this case the distribution of the critical points is always the same as the distribution of the zeros.

Example 9 For $P_n(z) = z^n - 1$ the zeros are the n-th roots of unity, but P_n' has all of its zeros at the origin. Thus, in this case the distribution of the zeros is the normalized arc measure λ on the unit circle C_1, but the distribution of the critical points is the Dirac mass at the origin.

The following example is a slight variation of the preceding one with a completely different conclusion.

Example 10 Let $Q_n(z) = z(z^{n-1} - 1)$. The distribution of the zeros is again the normalized arc measure λ on the unit circle C_1, but since $Q_n'(z) = nz^{n-1} - 1$, we obtain that λ is also the distribution of the critical points.

The main difference in between Examples 8 and 9 is that in the first one the complement of the support S of the limit measure μ is connected (since $S \subset \mathbf{R}$), while in the second example $S = C_1$, the complement of which is not connected. Indeed, in the case when the complement is connected, the distribution of the critical points always follows the distribution of the zeros (see [28, Theorem 1.1]).

Theorem 11 *If $\nu_{P_n} \to \mu$ and the support S of μ has connected complement, then $\nu_{P_n'} \to \mu$.*

Example 9 shows that this may not be the case if the complement of S is not connected. However, this example is very special: indeed, if $S = C_1$ but μ is not the normalized arc measure on C_1, then necessarily $\nu_{P_n'} \to \mu$. This is a special case of a general principle that we formulate now. If Γ is a Jordan curve (homeomorphic image of a circle) then its equilibrium measure μ_Γ is the unique probability measure on Γ for which the logarithmic potential

$$U^{\mu_\Gamma}(z) = \int \log \frac{1}{|z-t|} d\mu_\Gamma(t)$$

is constant on Γ. For example, (by symmetry) μ_{C_1} is the normalized arc measure λ on the unit circle C_1. Now it turns out that if $\nu_{P_n} \to \mu$, the support S of μ lies on a Jordan curve Γ and $\mu \neq \mu_\Gamma$, then the distribution of the critical points is again μ. For analytic Γ there are however, examples of polynomials P_n, $n = 1, 2, \ldots$, such that $\nu_{P_n} \to \mu_\Gamma$, but all limits ν of $\nu_{P_n'}$ have support lying strictly inside Γ (and so, in particular, $\nu \neq \mu_\Gamma$). Such examples are, however, very unstable, if we change a zero of P_n (by an amount $\geq 1/n^\gamma$) or delete or add (like in Example 10) a zero, then for the resulting polynomials the distribution of the critical points will be already μ_Γ. There are also Jordan curves (like curves with an inner angle $< \pi$ at some point) for which even the case $\mu = \mu_\Gamma$ is not an exception, i.e. on those curves the distribution of the critical points always agrees with the distribution of the zeros. For all these results see [28, Theorem 1.2].

Finally, let us discuss what happens in the general case, i.e. when the support S is not lying on a Jordan curve. The set $\mathbf{C} \setminus S$ has an unbounded component Ω and bounded components $\{G_j\}_{j=1}^J$ (their number may be infinite, finite or even zero, in which case we set $J = 0$). We define the inner boundary of S as the closure of the union of the boundaries of the connected components:

$$\partial_{\text{inner}} S = \overline{\cup_{j=1}^J \partial G_j},$$

while the outer boundary is the boundary $\partial\Omega$ of the unbounded component Ω of $\mathbf{C} \setminus S$. The inner and outer boundaries may not be disjoint, and together they give the boundary of S. With this notation we have the following theorem (see [28, Theorem 1.6]).

Theorem 12 *Suppose that $\nu_{P_n} \to \mu$, where μ is a unit measure with compact support S.*

- *If $\mu(\partial_{\mathrm{inner}} S) = 0$, then $\nu_{P'_n} \to \mu$.*
- *If G_j are the connected components of $\mathbf{C} \setminus S$ and if $O = \mathbf{C} \setminus \overline{\cup_j G_j}$, then for any weak*-limit ν of $\{\nu_{P'_n}\}$ we have*

$$\mu|_O = \nu|_O.$$

Note that if S has connected complement, then $O = \mathbf{C}$, so the second part of the theorem implies Theorem 11. Note also that the interior of S lies in O, hence in the interior of S the distribution of the critical points is the same as the distribution of the zeros. In particular, if the zeros are distributed according to an area-like measure μ, then the distribution of the critical points is μ (conjectured by B. Shapiro).

6 Generalizations, sharper forms, special cases

In some cases further restrictions on the location of the critical points can be given. We sample below a few such results which can be used in conjunction with the Gauss-Lucas theorem or with each other (when they are applicable).

Real polynomials

For real polynomials J. L. W. V. Jensen [10] stated the following theorem. Recall that if P is real, then its non-real roots can be paired into complex pairs $a_j \pm ib_j$, $b_j > 0$. For each such complex pair of roots let $D_j = \{z \mid |z - a_j| \le b_j\}$ be the disk over the segment connecting the zeros $a_j \pm ib_j$.

Theorem 13 (Jensen) *If P is real, then the non-real zeros of P' all lie in the union of the disks D_j.*

The following simple proof is from [5]. Let $z_{j\pm} = a_j \pm ib_j$ be a pair of complex conjugate roots and set $z = x + iy$. Simple computation shows that the imaginary part of

$$\frac{1}{z - z_{j+}} + \frac{1}{z - z_{j-}}$$

is

$$\frac{-2y[(x - a_j)^2 + y^2 - b_j^2]}{|z - z_{j+}|^2 |z - z_{j-}|^2},$$

so outside the disk D_j it is of opposite site to y. In a similar vein, for a real zero a_k of P

$$\Im \frac{1}{z-a_k} = \frac{-y}{|z-a_k|^2},$$

which is again of opposite site to y. Therefore, for z lying outside the real line (i.e. for $y \neq 0$) and of $\cup_j D_j$, the imaginary part of the sum in (1) is not zero, hence $P'(z) \neq 0$.

■

Circular domains

Let us start with J. H. Grace's theorem from [9].

Theorem 14 (Grace) *If z_1, z_2 are any two zeros of a polynomial P of degree n, then P' has a zero in the disk with center at $(z_1+z_2)/2$ and of radius $\frac{1}{2}|z_1 - z_2|\cot(\pi/n)$.*

By the Gauss-Lucas theorem if all zeros of a polynomial lie in a disk, then the same disk contains all the critical point. J. L. Walsh's [31] classical two-circle theorem discusses the case when the zeros lie in two disks.

Theorem 15 (Walsh) *Let D_1, D_2 be two disks with center at c_1, c_2 and of radius r_1, r_2, respectively. Let P be a polynomial of degree n with all its zeros in $D_1 \cup D_2$, say n_1 zeros lie in D_1 and n_2 zeros lie in D_2. Then P has all its critical points in $D_1 \cup D_2 \cup D_3$, where D_3 is the disk with center at $(n_1c_2 + n_2c_1)/n$ and of radius $(n_1r_2 + n_2r_1)/n$.*

Furthermore, if these three disks are pairwise disjoint, then D_1 contains $n_1 - 1$ critical points, D_2 contains $n_2 - 1$ critical points, and D_3 contains 1 critical point.

Next, suppose that $P(z) = \prod_{j=1}^m (z-z_j)^{k_j}$, where the z_j's are different, so the degree of P is $n = k_1 + \cdots + k_m$. The paper [4] by D. Dimitrov defines subdomains of the convex hull that contains the non-trivial critical points of P (i.e. those critical points that are different from every z_j). To describe his results, for each $1 \leq j \leq m$ choose a closed disk D_j that contains the points k_j/n and 1, and for $l \neq j$ set $D_{jl} = z_j + (z_l - z_j)D_j$. This is an affine transform of D_j that contains the point x_l and the point X_{jl} which divides the line segment $\overline{z_j z_l}$ in the ratio $k_j/(n-k_j)$. Finally, set

$$\Omega_j = \cup_{l \neq j} D_{jl}.$$

Theorem 16 (Dimitrov) *Every critical point of P different from z_j belongs to Ω_j. As a consequence, all non-trivial critical point belongs to $\cap_j \Omega_j$.*

In particular, we may take D_{jl} to be the disk with diameter $\overline{X_{jl}z_l}$ and we obtain the corollary that every critical point different from z_j belongs to the union (for $l \neq j$) of the disks with center at

$$\frac{n-k_j}{2n}z_j + \frac{n+k_j}{2n}z_l$$

and of radius

$$\frac{n-k_j}{2n}|z_l - z_j|. \tag{6}$$

As an immediate corollary it follows that no critical point different from z_j lies closer to z_j than

$$\frac{k_j}{n}\min_{l\neq j}|z_l - z_j|.$$

This is a quantitative manifestation of the fact (coming from the electrostatic interpretation of the critical points given in the first section) that no critical point different from z_j can lie too close to z_j.

The above construction is sharp in some sense, as is shown by

Example 17 Let $P(z) = z^3 - 1$ (or any power of $z^3 - 1$). Then z_1, z_2, z_3 are the third roots of unity, and $k_j/n = 1/3$ for all j. Simple geometry shows that if D_{12} is the disk with diameter $\overline{X_{12}z_2}$, where $X_{12} = (2z_1 + z_2)/3$ is the point (lying closer to z_1) that trisects the segment $\overline{z_1z_2}$, and D_{13} is formed similarly for the pair z_1, z_3, then D_{12} and D_{13} touch each other at the origin, which is only critical point of P. Thus, in this case the critical point lies on the boundaries of the disks described above, hence no smaller radii than what is given in (6) would work.

Strips

Let again the polynomial P have zeros $z_1, \ldots, z_n$. Consider a direction (a non-zero vector) w on the plane and draw a line trough each z_j parallel with w. These lines have the equations $ax + by + c_j = 0$ with the same a, b (that depend on w) and with possibly different c_j for different z_j. Consider the polynomial $Q(z) = \prod_{j=1}^n (z - c_j)$ with real zeros, and let $c_1', \ldots, c_{n-1}'$ be the (real) zeros of Q'. If $c_{\min}'$ and $c_{\max}'$ are the smallest and largest of them, then consider the lines $\ell_{\min}', \ell_{\max}'$ through them that are parallel with w. These form a closed strip that we denote by S_w. Formula (4) with $k = 1$ shows (since we may assume without loss of generality that the direction w is vertical) that all zeros of P' lie in S_w, which proves the following theorem of B. Z. Linfield [11] (see also [2]).

Theorem 18 (Lindfeld) *The critical points of P lie in $\cap_w S_w$.*

It can also be shown that the non-trivial critical points lie in the interior of the strip S_w.

Degrees three and four

A beautiful theorem of J. Siebeck [26] is the following.

Theorem 19 (Siebeck) *If P has degree three and its roots form the triangle $T = z_1z_2z_3$, then the critical points of P are at the foci of the only conic which is tangent to the sides of T at their midpoints.*

If P is of the form $(z-z_1)^{k_1}(z-z_2)_2^k(z-z_3)^{k_3}$, then the non-trivial critical points are at the foci of the conic that touches the sides of T and the points of tangency divide the corresponding sides of T in the ratio k_1/k_2, k_1/k_3 and k_2/k_3, respectively. See [14, Theorem 4.1].

There is a higher-order version (allowing any number of zeros and higher order algebraic curves) due to B. Z. Linfeld [11] (see also [2]).

Let us now consider a polynomial P of degree 4 with zeros z_1, z_2, z_3, z_4, when we assume that z_4 lies inside the triangle $T = z_1z_2z_3$. By connecting z_4 with z_1, z_2, z_3 we get a division of T into three triangles T_1, T_2, T_3. One is tempted to think that each of those triangles contains a critical point. A recent result of A. Rüdiger [23] claims that this is never the case.

Theorem 20 (Rüdiger) *The interior of at least one of T_1, T_2, T_3 is free from critical points.*

For the case when the degree of P is $n \geq 4$ but the convex hull of the zeros has fewer than n sides see [23].

7 The sector theorem

In this section we consider polynomials with nonnegative coefficients and the sectors

$$K_\theta = \{z \,|\, \operatorname{Arg} z \geq \theta\},$$

where we take the main branch of the argument. If $\theta \geq \pi/2$, then K_θ is convex, hence, by the Gauss-Lucas theorem, if K_θ contains all zeros of a polynomial P, then it contains all of its critical points, as well. It was proved by B. Sendov and H. Sendov [25] that the claim is actually true even if K_θ is not convex provided P has nonnegative coefficients.

Theorem 21 (Sendov and Sendov) *Let $0 < \theta < \pi$. If K_θ contains all zeros of a polynomial with nonnegative coefficients, then it contains all of its critical points.*

As a corollary we deduce the following. Let $\mathbf{C}_+$ be the closed upper half plane, and let $\mathcal{K}_+$ be the set of all closed convex sets $K_+ \subset \mathbf{C}_+$ which have the property that $K_+ \cap \mathbf{R} = [-a, b]$ with some $a, b \geq 0$ (possibly one or both $+\infty$), and K does not have a point in the half-plane $\{z \,|\, z < -a\}$. Denote the reflection of K_+ onto $\mathbf{R}$ by K_+^T, and set $\mathcal{K} = \{K_+ \cup K_+^T \,|\, K_+ \in \mathcal{K}_+\}$. Sets in $\mathcal{K}$ include (possibly non-convex) closed sectors with vertex at a point $c \in [0, \infty)$ and with axis of symmetry $(-\infty, c]$, or the cardioid $r = 1 - \cos\varphi$ (in polar form).

Corollary 22 *If $K \in \mathcal{K}$ and P is a polynomial with nonnegative coefficients which has all its zeros in K, then the same is true of P'.*

The proof of Theorem 21 in [25] is highly non-trivial, it involves the argument principle along with very careful zero and sign counting. We present a short proof along similar lines. We prove the claim in the following form.

Theorem 23 *Let $p(z) = a_0 + a_1 z + \cdots a_n z^n$, $n \geq 1$, $a_n \neq 0$, be a polynomial with nonnegative coefficients. If $0 < \theta < \pi$, and p has no zero in the sector*

$$S_\theta := \{z \,|\, 0 < \operatorname{Arg}(z) < \theta\},$$

then the same is true of p'.

Since both p and p' are real (hence their zeros are symmetric with respect to the real line) the two forms are clearly equivalent.

Proof. Considering instead of $p(z)$ the polynomial $p(z+\varepsilon)$ for some small $\varepsilon > 0$, we may assume that $a_j > 0$ for all $0 \leq j \leq n$ and that p and p' have no zeros on the boundary of S_θ.

The counterclockwise oriented boundary ∂K_R of $K_R := S_\theta \cap \{z \,|\, |z| < R\}$ consists of the segments $[0, R]$, $[Re^{i\theta}, 0]$ and the counterclockwise arc J_R on the circle $\{z \,|\, |z| = R\}$ that connects the points R and $Re^{i\theta}$. Since p has no zero in S_θ, we get from the argument principle that the total change of the argument of $p(z)$ on ∂K_R is 0. But the change of the argument over $[0, R]$ is 0 and over J_R is $n\theta + O(1/R)$, hence the change of the argument over $[Re^{i\theta}, 0]$ is $-n\theta + O(1/R)$. Upon letting $R \to \infty$ we can conclude that the total change of the argument of $p(te^{i\theta})$ along $t \in [0, \infty)$ is $n\theta$.

Let $f(t) := \arg(p(te^{i\theta}))$, where we choose that branch of the argument for which $f(0) = \arg(p(0)) = 0$. We claim that f increases on $[0, \infty)$. Indeed, suppose this is not the case. Then there are $0 < t_1 < t_2$ such that $f(t_2) < f(t_1)$. Since $f'(0) = \arg(a_1 e^{i\theta}) = \theta > 0$, then such a t_1, t_2 can be chosen with $f(t_2) > 0$. Let $\psi + k_0\pi \in (f(t_2), f(t_1))$ be a point such that $0 < \psi < \pi$, $k_0 \in \mathbf{N}$, and $j\theta - \psi$ is not an integer multiple of π for any integer j. Let $k \geq -1$ be the integral part of $(n\theta - \psi)/\pi$. f is a real valued continuous function on $[0, \infty)$ such that $f(0) = 0$ and $\lim_{t\to\infty} f(t) = n\theta$, hence its graph intersects each of the $(k+1)$ horizontal lines $y = s\pi + \psi$, $0 \leq s \leq k$, at least once. If $0 \leq k_0 \leq k$, then the graph intersects $y = k_0\pi + \psi$ at least three times by the choice of k_0 and ψ, so in this case $P(t) := \Im(e^{-i\psi} p(te^{i\theta}))$ has at least $k+3$ zeros on $(0, \infty)$. When $k = -1$ or $k_0 > k$, then we can make the same conclusion, since in these cases the graph of f intersects the line $y = k_0\pi + \psi$ at least twice. $P(t)$ is a real polynomial in $t \geq 0$ with coefficients $a_j \sin(j\theta - \psi)$, so, by Descartes' rule of sign, there are at least $k+3$ sign changes among the coefficients of P. But in between any two sign changes of the sequence $\{\sin(j\theta-\psi)\}_{0 \leq j \leq n}$ there is a zero of the function $\Im(e^{i(t\theta-\psi)})$, $t \geq 0$, so the curve $\{e^{i(t\theta-\psi)}\}$, $0 \leq t \leq n$, crosses the real axis at least $k+3 > (n\theta-\psi)/\pi - 1 + 3 > n\theta/\pi + 1$ times, which is not the case, since, as it is easy to see, the number of intersections is smaller than $n\theta/\pi + 1$. This contradiction proves the claim about the monotonicity of f.

Consider now the curve $\Gamma(t) := p(te^{i\theta})$, $t \in [0, \infty)$. We have shown that $f(t) = \arg(\Gamma(t))$ increases, so at any time t the curve Γ moves from the point $\Gamma(t)$ to the half plane that lies to the left of the (directed) half-line $\{s\Gamma(t) \mid s \geq 0\}$. Hence, the unit tangent vector to $\Gamma(t)$, i.e. $\Gamma'(t)/|\Gamma'(t)|$, is obtained from the direction of the position vector, i.e. from $\Gamma(t)/|\Gamma(t)|$, by a counterclockwise rotation with angle $\in [0, \pi]$. Thus, for any values of the arguments we always have

$$\arg(\Gamma(t)) \leq \arg(\Gamma'(t)) \leq \arg(\Gamma(t)) + \pi \tag{7}$$

mod 2π. But (7) is true at $t = 0$ by the choice of the branch of the argument function in f, hence (7) remains true (not just mod 2π!) for every $t \geq 0$. Furthermore, as $t \to \infty$ the two values $\arg(\Gamma(t))$ and $\arg(\Gamma'(t))$ agree ($= n\theta$) mod 2π, which is possible in view of (7) only if $\arg(\Gamma(t)) - \arg(\Gamma'(t)) \to 0$ as $t \to \infty$. But $\arg(\Gamma(0)) = 0$, $\arg(\Gamma'(0)) = \arg(a_1 e^{i\theta}) = \theta$, so it follows that the total change of the argument in $\Gamma'(t)$ over $[0, \infty)$ is θ less than the total change of the argument in Γ. Since this latter one is $n\theta$ by the first part of the proof, we obtain that the total change of the argument in Γ' over $[0, \infty)$ is $(n-1)\theta$. This is the same as the total change of the argument for $e^{-i\theta}\Gamma' = \{p'(te^{i\theta}) \mid t \geq 0\}$.

Thus, as $R \to \infty$, the change of the argument of $p'(z)$ over the segment $[Re^{i\theta}, 0]$ is $-(n-1)\theta + o(1)$, and since its change over $[0, R]$ is 0 and its change over J_R is $(n-1)\theta + O(1/R)$, we can conclude that for all large R the total change of the argument over the boundary ∂K_R (which is always an integer multiple of 2π) must be 0. Then the argument principle gives that p' has no zero in K_R for any R, i.e. no zero in the sector in S_θ.

■

Proof of Corollary 22. Suppose that $z_0 \notin K$ is a zero of P', and assume, for example, that $\Im z_0 \geq 0$. We have $K = K_+ \cup K_+^T$ where $K_+ \in \mathcal{K}_+$. The assumption on K^+ implies that

- every supporting line ℓ to K_+ with positive tangent either intersects $[0, \infty)$, or K_+ and $-\infty$ lie on different sides of ℓ,
- every supporting line with negative tangent intersects $[0, \infty)$.

Since $z_0 \notin K \supset K_+$, there is a supporting line ℓ to K_+ that separates z_0 and K_+. If this ℓ is horizontal, say it has equation $y = b$ for some $b \geq 0$, then $\Im z_0 > b$ while K lies in the lower half plane determined by ℓ. But this is impossible by the Gauss-Lucas theorem, so we may assume that ℓ is not horizontal. There are then two possibilities:

ℓ does not intersect $[0, \infty)$. In this case ℓ has positive tangent and K_+ and $-\infty$ and K_+ lie on different sides of ℓ. Then $\ell \cap \mathbf{C}_+$ and its reflection onto $\mathbf{R}$ determines a closed sector of opening angle $< \pi$ which contains K but which does not contain z_0, which is again impossible by the Gauss-Lucas theorem.

ℓ intersects $[0, \infty)$, say $\ell \cap [0, \infty) = c$. The half line $\ell \cap \mathbf{C}_+$ and its reflection onto $\mathbf{R}$ determines an open sector $S_{\theta,c}$ with vertex at c and with axis of symmetry (c, ∞) such that $S_{\theta,c}$ does not contain a zero of P, but contains z_0. If we write P in terms of powers of $z - c$ (i.e. write $z = (z - c) + c$ and expand $P(z)$), then the so obtained polynomial still has nonnegative coefficients, and by writing z instead of $z - c$ we may assume $c = 0$. But then $z_0 \in S_{\theta,0}$ contradicts Theorem 23, and this contradiction proves the corollary.

■

References

[1] N. G. de Bruijn and T. A. Springer, On the zeros of a polynomial and of its derivative, II. *Nederl. Akad. Wetensch.*, **50**(1947) 264–270, *Indag. Math.*, **9**(1947), 458–464.

[2] E. Casas-Alvero, Siebeck curves and two refinements of the Gauss-Lucas theorem. *Math. Scand.*, **111**(2012), 12–41.

[3] J. L. Daz-Barrero and J. J. Egozcue, A generalization of the Gauss-Lucas theorem. *Czechoslovak Math. J.*, **58**(133)(2008), 481–486.

[4] D. Dimitrov, A refinement of the Gauss-Lucas theorem. *Proc. Amer. Math. Soc.*, **126**(1998), 2065–2070.

[5] W. H. Echols, Note on the roots of the derivative of a polynomial. *Amer. Math. Monthly*, **27**(1920), 299–300.

[6] P. Erdős and I. Niven, On the roots of a polynomial and its derivative. *Bull. Amer. Math. Soc.*, **54**(1948), 184–190.

[7] C. F. Gauss, *Collected Works*. Leipzig, Teubner, 19001903, vol. 3, p. 112, and vol. 8, p. 32, and vol. 9, p. 187.

[8] T. Genchev and B. Sendov, A note on the theorem of Gauss for the distribution of the zeros of a polynomial on the complex plane. *Fiz.-Math. Sp.*, **1**(1958), 169-171. (Bulgarian).

[9] J. H. Grace, The zeros of a polynomial. *Proc. Cambridge Philosophical Soc.*, **11**(1902), 352–357.

[10] J. L. W. V. Jensen, Recherches sur la théorie dés équations. *Acta Mathematica*, **36**(1913), 181–195.

[11] B. Z. Linfield, On certain polar curves with their application to the location of the roots of the derivatives of a rational function. *Trans. Amer. Math. Soc.*, **27**(1920), 17–21.

[12] F. Lucas, Propriétés géométriques des fractions rationelles. *C. R. Acad. Sci. Paris*, **77**(1874), 431–433; **78**(1874), 140–144; **78**(1874), 180–183; **78**(1874), 271–274.

[13] S. M. Malamud, Inverse spectral problem for normal matrices and the Gauss-Lucas theorem. *Trans. Amer. Math. Soc.*, **357**(2005), 4043–4064.

[14] M. Marden, *Geometry of Polynomials.* 2nd ed., Math. Surveys 3, Amer. Math. Soc., Providence, RI 1966.

[15] N. Nikolov and B. Sendov, A converse of the Gauss-Lucas theorem. *Amer. Math. Monthly*, **121**(2014), 541–544.

[16] R. Pereira, Differentiators and the geometry of polynomials. *J. Math. Anal. Appl.*, **285**(2003), 336–348.

[17] Q. I. Rahman and G. Schmeisser, *Analytic Theory of Polynomials.* Clarendon Press, Nov 7, 2002.

[18] M. Ravichandran, Principal submatrices, restricted invertibility and a quantitative Gauss-Lucas theorem. *International Math. Res. Notices*, rny163 (2018), https://doi.org/10.1093/imrn/rny163

[19] T. Richards, Level curve configurations and conformal equivalence of meromorphic functions. *Comput. Methods Funct. Theory*, **15**(2015), 323–371.

[20] T. J. Richards, On approximate Gauss-Lucas theorems. arXiv1706.05410.

[21] T. Richards, Some Recent Results on the Geometry of Complex Polynomials: The Gauss–Lucas Theorem, Polynomial Lemniscates, Shape Analysis, and Conformal Equivalence. arXiv:1910.01159, to appear in *Complex Analysis and its Synergies.*

[22] T. J. Richards and S. Steinerberger, Leaky roots and stable Gauss-Lucas theorems. *Complex Var. Elliptic Equ.*, **64**(2019), 1898–1904.

[23] A. Rüdiger, Strengthening the Gauss-Lucas theorem for polynomials with zeros in the interior of the convex hull. arXiv:1405.0689

[24] T. Sheil-Small, *Complex Polynomials.* Cambridge University Press, 2002.

[25] B. Sendov and H. Sendov, On the zeros and critical points of polynomials with nonnegative coefficients: a nonconvex analogue of the Gauss-Lucas theorem. *Constr. Approx.*, **46**(2017), 305–317.

[26] J. Siebeck, Über eine neue analytische Behandlungweise der Brennpunkte. *J. Reine Angew. Math.*, **64**(1864), 175–182.

[27] V. Totik, The Gauss-Lucas theorem in an asymptotic sense. *Bull. London Math. Soc.*, **48**(2016), 848–854.

[28] V. Totik, Distribution of critical points of polynomials. *Trans. Amer. Math. Soc.*, **372**(2019), 2407–2428.

[29] V. Totik, A quantitative Gauss-Lucas theorem (manuscript).

[30] E. B. Van Vleck, On the location of roots of polynomials and entire functions. *Bull. Amer. Math. Soc.*, **35**(1929), 643–683.

[31] J. L. Walsh, On the location of the roots of the derivative of a polynomial. *C. R. Congr. Internat. des Mathématiciens, Strasbourg*, 1920, 339–342.

MTA-SZTE Analysis and Stochastics Research Group
Bolyai Institute, University of Szeged
Szeged, Aradi v. tere 1, 6720, Hungary

totik@math.u-szeged.hu