

A STABILITY VERSION OF THE GAUSS-LUCAS THEOREM AND APPLICATIONS

STEFAN STEINERBERGER

ABSTRACT. Let $p : \mathbb{C} \rightarrow \mathbb{C}$ be a polynomial. The Gauss-Lucas theorem states that its critical points, $p'(z) = 0$, are contained in the convex hull of its roots. We prove a stability version whose simplest form is as follows: suppose p has $n + m$ roots where n are inside the unit disk,

$$\max_{1 \leq i \leq n} |a_i| \leq 1, \quad \text{and } m \text{ are outside} \quad \min_{n+1 \leq i \leq n+m} |a_i| \geq d > 1 + \frac{2m}{n},$$

then p' has $n - 1$ roots inside the unit disk and m roots at distance at least $(dn - m)/(n + m) > 1$ from the origin and the involved constants are sharp. We also discuss a pairing result: in the setting above, for n sufficiently large each of the m roots has a critical point at distance $\sim n^{-1}$.

1. INTRODUCTION AND RESULTS

1.1. Introduction. The Gauss-Lucas Theorem, first stated by Gauss [7] in 1836 and first proved by Lucas [12] in 1879, states that if $p : \mathbb{C} \rightarrow \mathbb{C}$ is a polynomial of degree n , then the $n - 1$ zeroes of p' lie inside the convex hull of the n zeros of p . This has been refined in various ways [2, 3, 4, 5, 6, 9, 11, 13, 14, 16, 19, 21, 22, 23, 24, 25]. It was recently established by Totik [26] that, for sequences of polynomials p_n with $\deg(p_n) \rightarrow \infty$, that if $n - o(n)$ roots of p lie inside a convex domain K , then any fixed neighborhood of K contains $n - o(n)$ roots of p' . We prove a sharp, non-asymptotic result in the same spirit.

Theorem 1. *Let $p : \mathbb{C} \rightarrow \mathbb{C}$ be a polynomial having n roots a_1, \dots, a_n inside the unit disk and m roots a_{n+1}, \dots, a_{n+m} outside. If the roots outside are bounded away from the disk*

$$\min_{n+1 \leq i \leq n+m} |a_i| \geq d > 1 + \frac{2m}{n},$$

then p' has $n - 1$ roots inside the unit disk and m roots of modulus at least $(dn - m)/(n + m) > 1$.

Theorem 1 is sharp: consider $p(x) = (x + 1)^n(x - d)^m$ for some real $d > 1$. The derivative p' has $n - 1$ roots in -1 , $m - 1$ roots in d and one root r in

$$r = \frac{dn - m}{n + m} \quad \text{which requires} \quad d > 1 + \frac{2m}{n} \quad \text{to be outside the unit disk.}$$



FIGURE 1. Roots outside create critical points outside the disk.

2010 *Mathematics Subject Classification.* 26C10, 30C15.

Key words and phrases. Gauss-Lucas theorem, polynomial dynamics, roots, zeroes, derivatives, critical points.

This work is supported by the NSF (DMS-1763179) and the Alfred P. Sloan Foundation.

The case $m = 1$ is essentially known, though phrased in a somewhat different language, and due to Rahman [20] (see also Aziz [1] for higher derivatives). We now formulate a version for general convex domains, $d(z, K)$ denotes the distance from a point $z \in \mathbb{C}$ to a set $K \subset \mathbb{C}$.

Theorem 2 (Stable Gauss-Lucas). *Let $p : \mathbb{C} \rightarrow \mathbb{C}$ be a polynomial having n roots a_1, \dots, a_n inside the convex domain $K \subset \mathbb{C}$ and m roots a_{n+1}, \dots, a_{n+m} outside. If the roots outside satisfy*

$$\min_{n+1 \leq i \leq n+m} d(a_i, K) \geq 2 \operatorname{diam}(K) \sqrt{\frac{m^2}{n^2} + \frac{m}{n}},$$

then p' has $n - 1$ roots close to K in the sense that

$$d(z, K) \leq \frac{\operatorname{diam}(K) \sqrt{m}}{\sqrt{m+n}}$$

and m other roots satisfying $d(z, K) \geq \operatorname{diam}(K) \sqrt{m/(n+m)}$.

We note that the result has a different scaling than Theorem 1 (the scaling in Theorem 1 is a consequence of the boundary of the unit disk having curvature bounded from below). The proof has various degrees of freedom (how to set which parameter etc.) and it is possible to obtain a variety of other results of a similar flavor with the same approach.

1.2. Pairings of Roots and Critical Points. There is recent renewed interest in the interplay between roots of a polynomial and the location of its critical points in the random setting (see e.g. [8, 10, 18]). Kabluchko [10] has shown that if a random polynomial is constructed by picking its n roots from some probability measure μ , then the roots of the derivative converge to the same measure μ as $n \rightarrow \infty$. The following recent result is due to O'Rourke & Williams [15]: suppose μ is a compactly supported probability measure, z_1, \dots, z_n are randomly drawn from μ and $\xi \in \mathbb{C}$ is a deterministic point outside the support of μ . If p denotes the random polynomial

$$p(z) = (z - \xi) \prod_{k=1}^n (z - z_k),$$

then p' has a root at distance $\sim n^{-1}$ from ξ with high likelihood. We give a deterministic variant.

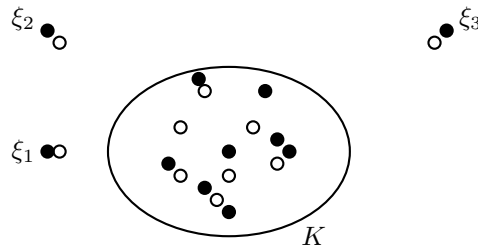


FIGURE 2. Roots outside (filled) have critical points (empty) nearby.

Theorem 3 (Pairing of Roots and Critical Points). *Let $K \subset \mathbb{C}$ be a closed convex domain and $\xi_1, \dots, \xi_m \in \mathbb{C} \setminus K$. Then there exists $n_0 \in \mathbb{N}$ and $c > 0$, both depending on m , $\operatorname{diam}(K)$ and $\min_{1 \leq i \leq m} \min_{z \in K} \|\xi_i - z\|$ such that for all $n > n_0$ and all $z_1, \dots, z_n \in K$, the polynomial*

$$p(z) = \left(\prod_{\ell=1}^m (z - \xi_\ell) \right) \prod_{k=1}^n (z - z_k)$$

has exactly m critical points outside of K and all of them are at distance $\leq c/n$ from $\{\xi_1, \dots, \xi_m\}$. Conversely, for each ξ_i , there is a critical point in a c/n -neighborhood.

We observe that this is somewhat different from the version of O'Rourke & Williams [15]: it is completely deterministic but also requires that the ξ_i are not contained inside the convex hull of the support of μ . It turns out that this condition (or some condition in that direction) is necessary in the deterministic setting (as was already observed in [15]): let K to be an annulus centered at the origin containing the boundary of the unit disk, let $\xi = 0$ and consider the polynomial

$$p(z) = (z - \xi)(z^n - 1) = z \prod_{k=1}^n \left(z - e^{2\pi i k/n} \right).$$

For every $\varepsilon > 0$ and n sufficiently large (depending on ε), all the critical points are contained in the annulus $\{z \in \mathbb{C} : 1 - \varepsilon \leq |z| \leq 1\}$ and none of them are particularly close to $\xi = 0$. The proof of Theorem 3 is fairly explicit and the constant c could be made explicit in terms of everything else if one so desired. The proof also yields the other direction as a byproduct: the critical points either coincide with an ξ_i (which can happen in the case of multiplicity) or are indeed at distance $\gtrsim c_2/n$ away from ξ_i for a constant c_2 (depending on the same things as c).

2. PROOFS

2.1. Proof of Theorem 1.

Proof. Let $p : \mathbb{C} \rightarrow \mathbb{C}$ be a polynomial of degree $n + m$, having n roots $\{a_1, \dots, a_{n+m}\}$ where $|a_i| \leq 1$ for all $1 \leq i \leq n$ and

$$d := \min \{|a_{n+1}|, |a_{n+2}|, |a_{n+3}|, \dots, |a_{n+m}|\} \geq 1.$$

For our argument it is not important whether the roots are distinct or occur with multiplicity. The derivative p' has $n + m - 1$ roots whose location is determined by the logarithmic derivative

$$\frac{p'(z)}{p(z)} = \sum_{k=1}^{n+m} \frac{1}{z - a_k}.$$

We start by obtaining a lower bound on the size of the quantity outside the unit disk caused by the terms contained inside the unit disk. An elementary inequality for real $x > 1$ and complex $|y| \leq 1$ is given by

$$\operatorname{Re} \frac{1}{x - re^{it}} \geq \frac{1}{x + 1} \quad \text{with equality for } (r, t) = (1, \pi).$$

This implies, for $z = re^{it}$ with $r > 1$,

$$\left| \sum_{k=1}^n \frac{1}{z - a_k} \right| = \left| \frac{1}{e^{it}} \sum_{k=1}^n \frac{1}{ze^{-it} - a_k e^{-it}} \right| \geq \operatorname{Re} \sum_{k=1}^n \frac{1}{ze^{-it} - a_k e^{-it}} \geq \frac{n}{|z| + 1}.$$

We now estimate the size of the remaining term for $|z| < d$. Clearly,

$$\left| \sum_{k=n+1}^{n+m} \frac{1}{z - a_k} \right| \leq \sum_{k=n+1}^{n+m} \frac{1}{|z - a_k|} \leq \frac{m}{d - |z|}.$$

In the location of a new root outside the unit disk, the electrostatic forces add up to 0 and we therefore have to have

$$\frac{n}{|z| + 1} \leq \left| \sum_{k=1}^n \frac{1}{z - a_k} \right| = \left| \sum_{k=n+1}^{n+m} \frac{1}{z - a_k} \right| \leq \frac{m}{d - |z|}$$

which then simplifies to

$$|z| \geq \frac{dn - m}{n + m}.$$

It remains to show that the derivative still has m roots outside the unit disk (which, by the preceding argument, are then necessarily at least a controlled distance away from the unit disk).

This is done by showing that p' has $n - 1$ roots inside the unit disk. We introduce a polynomial collecting all roots inside the unit disk

$$q(x) = \prod_{k=1}^n (z - a_k)$$

and observe that critical points of p are zeros of the function

$$h(z) = q'(z) + q(z) \sum_{k=n+1}^{n+m} \frac{1}{z - a_k}.$$

h is holomorphic in a neighborhood of the unit disk. The argument above shows that

$$|q'(z)| \geq \frac{n}{|z| + 1} |q(z)| \quad \text{for } z \text{ outside the unit disk.}$$

We now apply Rouché's theorem on the boundary of the unit disk (or a slightly larger disk if there happens to a root $|a_i| = 1$). Since

$$d > 1 + \frac{2m}{n} \quad \text{we have} \quad \frac{n}{|z| + 1} > \frac{m}{|d| - |z|}$$

for $|z| = 1$ and thus

$$\left| q(z) \sum_{k=n+1}^{n+m} \frac{1}{z - a_k} \right| < |q'(z)|.$$

This shows that the number of roots of p' inside the unit disk, which is the number of roots of h is exactly the same as the number of roots of q' which, by the Gauss-Lucas theorem, is $n - 1$. \square

2.2. Proof of Theorem 2.

Proof. The structure of the argument is completely identical to that of Theorem 1, however, some of the computational aspects change. As before we assume that $a_1, \dots, a_n \in K$ and abbreviate

$$d := \min \{d(a_{n+1}, K), d(a_{n+2}, K), d(a_{n+3}, K), \dots, d(a_{n+m}, K)\}.$$

Let $z \in \mathbb{C} \setminus K$, then we first require a lower bound on

$$\left| \sum_{k=1}^n \frac{1}{z - a_k} \right|.$$

By rotational and translational variance, we can again assume that $z \in \mathbb{R}$, that $(0, 0) \in K$ is the closest point in K and

$$z > \sup_{y \in K} \operatorname{Re} k = 0.$$

A simple computation shows that for all $k \in K$

$$\operatorname{Re} \frac{1}{z - k} = \operatorname{Re} \frac{1}{z - k_1 - ik_2} = \frac{z - k_1}{(z - k_1)^2 + k_2^2}.$$

We know that $k_1 \leq 0$ and that $k_1^2 + k_2^2 \leq \operatorname{diam}(K)^2$. An explicit optimization yields that

$$\begin{aligned} \frac{z - k_1}{(z - k_1)^2 + k_2^2} &\geq \begin{cases} \frac{d(z, K)}{d(z, K)^2 + \operatorname{diam}(K)^2} & \text{if } \operatorname{diam}(K) \geq d(z, K) \\ \frac{1}{d(z, K) + \operatorname{diam}(K)} & \text{if } \operatorname{diam}(K) \leq d(z, K) \end{cases} \\ &\geq \frac{d(z, K)}{d(z, K)^2 + \operatorname{diam}(K)^2}. \end{aligned}$$

Altogether, this implies

$$\left| \sum_{k=1}^n \frac{1}{z - a_k} \right| \geq \frac{n \cdot d(z, K)}{d(z, K)^2 + \operatorname{diam}(K)^2} \quad \text{for } z \in \mathbb{C} \setminus K.$$

The upper bound on the effect from outside K is unchanged

$$\left| \sum_{k=n+1}^{n+m} \frac{1}{z - a_k} \right| \leq \sum_{k=n+1}^{n+m} \frac{1}{|z - a_k|} \leq \frac{m}{d - d(z, K)}.$$

This shows that the argument in the proof of Theorem 1 is applicable as soon as

$$\frac{n \cdot d(z, K)}{d(z, K)^2 + \text{diam}(K)^2} > \frac{m}{d - d(z, K)}$$

This is satisfied as soon as

$$d > d(z, K) \left(1 + \frac{m}{n} \right) + \frac{\text{diam}(K)^2}{d(z, K)} \frac{m}{n}.$$

Minimizing this quantity shows in the variable $d(z, K)$ shows that we would like to apply it for

$$d(z, K) = \frac{\text{diam}(K) \sqrt{m}}{\sqrt{m+n}}$$

which ends up requiring

$$d > 2 \text{diam}(K) \sqrt{\frac{m^2}{n^2} + \frac{m}{n}}.$$

□

2.3. Proof of Theorem 3.

Proof. Theorem 2 immediately implies that for $\{\xi_1, \dots, \xi_m\}$ and n sufficiently large, there are exactly m critical points outside of K . It remains to understand their location. Denoting the roots inside K by z_1, \dots, z_n , we obtain that any critical point z satisfies

$$\left| \sum_{\ell=1}^n \frac{1}{z - z_\ell} \right| = \left| \sum_{\ell=1}^m \frac{1}{\xi_\ell - z} \right| \leq \frac{m}{\min_{1 \leq \ell \leq m} |z - \xi_\ell|}.$$

We know from the proof of Theorem 2 that the left-hand side grows like

$$\left| \sum_{\ell=1}^n \frac{1}{z - z_\ell} \right| \geq \frac{n \cdot d(z, K)}{d(z, K)^2 + \text{diam}(K)^2} \quad \text{for } z \in \mathbb{C} \setminus K.$$

This requires the right-hand side to be in a $\sim n^{-1}$ neighborhood of $\{\xi_1, \dots, \xi_m\}$. It also shows that the constant c depend on m , the distance of $\{\xi_1, \dots, \xi_m\}$ to K and the diameter of K . It remains to show that a c/n -neighborhood of ξ_1 necessarily contains a critical point. Critical points satisfy

$$\frac{\text{mult}(\xi_1)}{z - \xi_1} = \sum_{k=2}^m \frac{1}{\xi_k - z} + \sum_{k=1}^n \frac{1}{z_k - z},$$

where $\text{mult}(\xi_1)$ denotes the number of times ξ_1 appears in the list and thus w.l.o.g. $\xi_1 \neq \xi_i$ for all $2 \leq i \leq m$. This equation can be rewritten as

$$(z - \xi_1) - \text{mult}(\xi_1) \left(\sum_{k=2}^m \frac{1}{\xi_k - z} + \sum_{k=1}^n \frac{1}{z_k - z} \right)^{-1} = 0. \quad (\diamond)$$

By construction, the first term in parentheses is bounded by a constant in a sufficiently small neighborhood of ξ_1 , i.e. there exists $\varepsilon > 0$ such that for all $|z - \xi_1| \leq \varepsilon$ we have

$$\left| \sum_{k=2}^m \frac{1}{\xi_k - z} \right| \leq C.$$

The second term is growing linearly in n . For n sufficiently large

$$\frac{c_1}{n} \leq \left| \text{mult}(\xi_1) \left(\sum_{k=2}^m \frac{1}{\xi_k - z} + \sum_{k=1}^n \frac{1}{z_k - z} \right)^{-1} \right| \leq \frac{c_2}{n}$$

in a sufficiently small neighborhood of ξ_1 . Applying Rouché's theorem to (\diamond) on the boundary of the disk $\{z : |z - \xi_1| = 2c_2n^{-1}\}$ shows the existence of a critical point in a $2c_2n^{-1}$ neighborhood. Applying it in the other direction on the boundary of the disk $\{z : |z - \xi_1| = c_1n^{-1}/2\}$ shows that the scaling is optimal, i.e. that the critical point is actually at distance $\sim n^{-1}$ from ξ_1 . \square

REFERENCES

- [1] A. Aziz, On the Location of Critical Points of Polynomials, J. Aust. Math. Soc. 36 (1984), 4–11.
- [2] H. E. Bray. On the Zeros of a Polynomial and of Its Derivative. Amer. J. Math., 53(4):864–872, 1931.
- [3] N. G. de Bruijn. On the zeros of a polynomial and of its derivative. Nederl. Akad. Wetensch., Proc., 49:1037–1044 = Indagationes Math. 8, 635–642 (1946), 1946.
- [4] N. G. de Bruijn and T. A. Springer. On the zeros of a polynomial and of its derivative. II. Nederl. Akad. Wetensch., Proc., 50:264–270=Indagationes Math. 9, 458–464 (1947), 1947.
- [5] B. Curgus and V. Mascioni, A Contraction of the Lucas Polygon, Proc.Amer.Math.Soc. 132, 2004, 2973–2981.
- [6] D. Dimitrov, A Refinement of the Gauss-Lucas Theorem, Proc. Amer. Math. Soc. 126, 1998, 2065–2070.
- [7] C.F. Gauss: Werke, Band 3, Göttingen 1866, S. 120:112
- [8] B. Hanin, Pairing of zeros and critical points for random polynomials, Ann. Inst. H. Poincaré, Probab. Statist. 53(2017), 1498–1511.
- [9] A. Joyal. On the zeros of a polynomial and its derivative. J. Math. Anal. Appl., 26:315–317, 1969.
- [10] Z. Kabluchko., Critical Points of Random Polynomials with Independent and Identically Distributed Roots. Proc. AMS. 143(2), 2015, p. 695–702.
- [11] D. Kalman, An Elementary Proof of Marden's Theorem, Amer. Math. Monthly 115: 330–338 (2008).
- [12] F. Lucas: Sur une application de la Mécanique rationnelle à la théorie des équations. in: Comptes Rendus de l'Académie des Sciences (89), Paris 1979, S. 224–226
- [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, A note on the zeroes of the sections of a partial fraction, Bulletin of the American Mathematical Society 51 (1945): 935–940.
- [15] S. O'Rourke and N. Williams, Pairing between zeros and critical points of random polynomials with independent roots, arXiv:1610.06248, to appear in Trans. Amer. Math. Soc.
- [16] P. Pawlowski, On the Zeros of a Polynomial and its Derivatives, Trans.Amer.Math.Soc.350 (1998), p.4461–4472.
- [17] Pelckmans, K., De Brabanter, J., Suykens, J., and De Moor, B. Convex clustering shrinkage. PASCAL Workshop on Statistics and Optimization of Clustering Workshop, London, (2005).
- [18] R. Pemantle, and I. Rivlin. The distribution of the zeroes of the derivative of a random polynomial. Advances in Combinatorics. Springer 2013. Editors: Kotsireas, Ilias S. and Zima, Eugene V. pp. 259–273.
- [19] R. Pereira, Differentiators and the geometry of polynomials, J. Math. Anal. Appl., 285 (2003), 336–348.
- [20] Q. I. Rahman, On the zeros of a polynomial and its derivative, Pacific J. Math. 41 (1972), 525–528.
- [21] M. Ravichandran, Principal submatrices, restricted invertibility and a quantitative Gauss-Lucas theorem, arXiv:1609.04187
- [22] T. Richards, On approximate Gauss-Lucas theorems, arXiv:1706.05410.
- [23] J.Siebeck, Über eine neue analytische Behandlungsweise der Brennpunkte, Journal für die reine und angewandte Mathematik, 64: 175–182, (1864)
- [24] G. Schmeisser, Majorization of the Critical Points of a Polynomial by Its Zeros, Computational Methods and Function Theory 2004, p. 95–103.
- [25] W. Specht: Eine Bemerkung zum Satze von Gauss-Lucas, in: Jahresbericht der Deutschen Mathematiker-Vereinigung (62), 1959, S. 85–92.
- [26] V. Totik, The Gauss-Lucas theorem in an asymptotic sense, Bull. London Math. Soc. 48, 2016, p. 848–854.
- [27] J. L. Walsh, The location of critical points of analytic and harmonic functions, Amer. Math. Soc. Colloq. Publ., Vol. 34, 1950.

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CT 06511, USA
E-mail address: stefan.steinerberger@yale.edu