The Gauss-Lucas theorem in an asymptotic sense*

Vilmos Totik†

July 6, 2016

Abstract

According to the Gauss-Lucas theorem, if all zeros of a polynomial lie in a convex set K, then all zeros of its derivative also lie in K. In this paper it is shown that if almost all zeros of polynomials lie in a convex set K, then almost all zeros of their derivatives lie in any fixed neighborhood of K.

1 Introduction

The Gauss-Lucas theorem [1] says that the zeros of the derivative of a polynomial lie in the convex hull of the zeros of the polynomial itself. In particular, if all zeros of a polynomial p_n lie in a convex set K, then all zeros of p_n' also lie in K. This is no longer true if one zero may lie outside K, for then K may not contain any zero of the derivative. Indeed, if $z_1, \ldots, z_{n-1}$ are distinct points in $[0, 1]$, then the polynomial $q_n(z) = (z - i)\prod_1^{n-1}(z - z_i)$ have all of its zeros in $[0, 1]$ with one exception, but q_n' have all its zeros outside $[0, 1]$. Strict convexity of the boundary would not help, either, for example, if K is the closed unit disk and T is a linear transformation that maps 1 to 1 and 0 to e^{ai} with some small $a > 0$, then the polynomial $p_n(z) = q_n(T^{-1}(z))$ with the previous q_n have all its zeros on the segment connecting the points 1 and e^{ia}, but for sufficiently small $a > 0$ the zeros of p_n' lie outside the unit disk.

In this note we prove that, contrary to such counterexamples, the Gauss-Lucas theorem holds in an asymptotic sense even if some of the zeros of the polynomial lie outside K. This may be convenient in applications, when one does not know that every single zero of p_n lies in K.

Let $\{p_n\}$ be polynomials of degree $n = 1, 2, \ldots$. We say that p_n have almost all of their zeros on K if p_n have $o(n)$ zeros outside K. Equivalently, if μ_n denotes the counting measure on the zeros of p_n, then $\mu_n(K)/n \to 1$ as $n \to \infty$.

*AMS Classification: 26C10, 31A15; Key words: zeros of polynomials, Gauss-Lucas theorem, potential theory

†Supported by ERC Advanced Grant No. 267055

Theorem 1 *If p_n, $n = 1, 2, \ldots$, have almost all of their zeros on the compact convex set K, then for every $\varepsilon > 0$ the derivatives p_n' have almost all of their zeros on K_ε, where K_ε is the ε-neighborhood of K.*

The examples discussed before show that in the claim it is necessary to consider K_ε, i.e. a slightly larger set then the original one.

The proof of the Gauss-Lucas theorem is very simple: if $z_1, \ldots, z_n$ are the zeros of the polynomial and z lies outside the convex hull of them, then there is a line ℓ that separates z from all z_j, and without loss of generality we may assume this line ℓ to be the imaginary axis and, say, $\Re z > 0$. But then it immediately follows that

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

cannot be zero, for all terms on the right have positive real part. Based on this elementary argument one would expect that Theorem 1 has an equally simple proof, but a more careful examination of the problem reveals that such a simple argument may not be available. The proof we give uses potential theory. At the end of the paper we sketch a short proof, based on a theorem of Malamud and Pereira, which works in the special case when all zeros lie in a fixed compact set.

Let us also mention that one cannot hope for an extension of Theorem 1 in the sense that if K contains at least αn of the zeros of p_n, then K_ε contains at least αn (or any fixed portion) of the zeros of p_n'. Indeed, $p_n(z) = z^n - 1$ has at least one third of its zeros in the rectangle $K = [1/4, 1] \times [-1, 1]$, but p_n' has no zero in $K_{1/8}$ whatsoever.

Acknowledgement. The author thanks Boris Shapiro for stimulating discussions. In particular, he brought the problem to the author's attention, and he formulated Theorem 1 as a conjecture.

2 Proof of Theorem 1

We shall use some basic facts from logarithmic potential theory, see for example the books [4] or [5] for the general theory.

Without loss of generality we may assume that p_n has leading coefficient 1, and that $K \subset B_{1/4}$, where B_r is the open disk about the origin of radius r. Let S be the ring $B_{1/2} \setminus K$.

Let μ_n be the zero counting measure of p_n, and ν_n the zero counting measure of p_n'. Suppose to the contrary that the claim is not true, and there is an $\varepsilon > 0$ and an $\alpha < 1$ such that for infinitely many n, say for $n \in \mathcal{N}$, we have $\nu_n(K_\varepsilon)/n < \alpha$. We shall get a contradiction.

Let $\mathcal{N}_1 \subset \mathcal{N}$ be a subsequence along which $\mu_n/n \to \mu$, $\nu_n/n \to \nu$ in the weak* topology on the closed Riemannian sphere. Then μ is supported on K,

$\mu(K)=1$, and $\nu(K)\le\alpha$. Below we show that, on the other hand, $\nu(K)=1$, and that will constitute the required contradiction.

In what follows we shall denote by m_2 the two dimensional Lebesgue-measure on the complex plane.

I. Claim: *There is a subsequence $\mathcal{N}_2\subset\mathcal{N}_1$ such that for m_2-almost all $z\in S$ we have*

$$\lim_{n\to\infty,\ n\in\mathcal{N}_2}\frac{p_n'(z)}{np_n(z)}=\int\frac{1}{z-t}d\mu(t). \tag{1}$$

Indeed, $\mu_n=\mu_n\big|_K+\mu_n\big|_{\mathbf{C}\setminus K}$, and since $\mu_n\big|_{\mathbf{C}\setminus K}(\mathbf{C})=o(n)$ by assumption, we have $\mu_n\big|_{\mathbf{C}\setminus K}/n\to 0$, in the weak* topology. Since $\mu_n/n\to\mu$ also in the weak* topology, we can conclude that $\mu_n\big|_K/n\to\mu$ in the weak* topology. Therefore, for any $z\in S$ we have

$$\lim_{n\to\infty,\ n\in\mathcal{N}_1}\frac{1}{n}\int\frac{1}{z-t}d\mu_n\big|_K(t)=\int\frac{1}{z-t}d\mu(t). \tag{2}$$

Since

$$\frac{1}{n}\int\frac{1}{z-t}d\mu_n(t)=\frac{p_n'(z)}{np_n(z)},$$

it is left to prove that along some subsequence $\mathcal{N}_2\subset\mathcal{N}_1$ we have

$$\lim_{n\to\infty,\ n\in\mathcal{N}_2}\frac{1}{n}\int\frac{1}{z-t}d\mu_n\big|_{\mathbf{C}\setminus K}(t)=0 \tag{3}$$

for m_2-almost all $z\in S$. But that is clear: since

$$\int_S\frac{1}{|z-t|}dm_2(t)\le C,\qquad z\in\mathbf{C},$$

with some constant C that depends only on S, we have

$$\int_S\left(\frac{1}{n}\int\frac{1}{|z-t|}d\mu_n\big|_{\mathbf{C}\setminus K}(t)\right)dm_2(z)\le C\frac{\mu_n(\mathbf{C}\setminus K)}{n}\to 0,$$

which implies that a subsequence of the function in the brackets in the integrand on the left tends to 0 for m_2-almost all $z\in S$, and this is stronger than (3).

■

II. Claim: *The integral on the right of* (1) *is non-zero in S.* Indeed, let $z\in S$. Then z and K can be separated by a line, and without loss of generality we may assume that this line is the $\Re z=a$ line with some $a\in\mathbf{R}$. Then $\Re z>a$, while for all $t\in K$ we have $\Re t<a$ (or vice versa), so $\Re(z-t)>0$ for all $t\in K$, which implies $\Re(1/(z-t)) > 0$ for all such t. Since μ is supported on K, we can conclude that

$$\Re \int \frac{1}{z-t} d\mu(t) = \int \left(\Re \frac{1}{z-t} \right) d\mu(t) > 0,$$

which proves the claim.

■

III. Claim: *For m_2-almost all $z \in S$ we have*

$$\lim_{n\to\infty,\ n\in\mathcal{N}_2} \frac{1}{n} \log \frac{|p_n'(z)|}{|p_n(z)|} = 0.$$

This is an immediate consequence of Claims I and II because $\log n/n \to 0$.

■

Let

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

denote the logarithmic potential of a measure ρ with compact support.

Since

$$\frac{1}{n} \log \frac{|p_n'(z)|}{|p_n(z)|} = \frac{1}{n} U^{\mu_n}(z) - \frac{1}{n} U^{\nu_n}(z),$$

we get that along the subsequence $\mathcal{N}_2$

$$\frac{1}{n} U^{\mu_n}(z) - \frac{1}{n} U^{\nu_n}(z) \to 0 \tag{4}$$

for m_2-almost all $z \in S$.

IV. Claim. *There is a subsequence $\mathcal{N}_3 \subset \mathcal{N}_2$ and a sequence $\{a_n\}$ of constants such that for m_2-almost all $z \in S$*

$$\lim_{n\to\infty,\ n\in\mathcal{N}_3} \left(\frac{1}{n} U^{\mu_n}(z) - a_n \right) = U^\mu(z). \tag{5}$$

We write $\mu_n = \mu_n^1 + \mu_n^2$, where μ_n^2 is the restriction of μ_n to the exterior of $\overline{B_{1/2}}$ (and hence μ_n^1 is the restriction of μ_n to $\overline{B_{1/2}}$). Let μ_n^3 be the balayage of μ_n^2 out of $\mathbf{C} \setminus \overline{B_{1/2}}$ (see e.g. section II.3 in [5] for the concept of balayage). Then μ_n^3 is a measure on $\partial B_{1/2}$ such that it has the same total mass as μ_n^2, and with some constant c_n we have

$$U^{\mu_n^2}(z) = U^{\mu_n^3}(z) + c_n, \qquad z \in \overline{B_{1/2}}.$$

Since the total mass of μ_n^3/n (which is the same as the total mass of μ_n^2/n) tends to 0, and this measure lies on the circle $|z| = 1/2$, it follows that

$$\frac{1}{n}U^{\mu_n^2}(z) - \frac{c_n}{n} = \frac{1}{n}U^{\mu_n^3}(z) \to 0, \qquad z \in B_{1/2}.$$

On the other hand, in the proof of claim I we have seen that with $\mu_n^0 := \mu_n|_K$ we have $\frac{1}{n}\mu_n^0 \to \mu$ in the weak* topology, which implies that

$$\frac{1}{n}U^{\mu_n^0}(z) \to U^{\mu}(z), \qquad z \in S.$$

Since $\mu_n = \mu_n^0 + (\mu_n^1 - \mu_n^0) + \mu_n^2$, it is left to prove that along some subsequence of $\mathcal{N}_3$ of $\mathcal{N}_2$ we have

$$\frac{1}{n}U^{\mu_n^1 - \mu_n^0}(z) \to 0 \tag{6}$$

for m_2-almost all $z \in S$.

The measure $\mu_n^1 - \mu_n^0$ is the restriction of μ_n to the set $\overline{B_{1/2}} \setminus K$, say $\mu_n^1 - \mu_n^0 = \sum_{k=1}^{m_n} \delta_{z_k^n}$, where, by assumption, $m_n/n \to 0$. Note that

$$h_n(z) := \frac{1}{n}U^{\mu_n^1 - \mu_n^0}(z) = \frac{1}{n}\int \log\frac{1}{|z-t|}d(\mu_n^1 - \mu_n^0)(t) = \frac{1}{n}\sum_{k=1}^{m_n}\log\frac{1}{|z - z_k|} \geq 0$$

on S because $z, z_k^n \in B_{1/2}$, and hence $|z - z_k^n| < 1$. Now with some $\varepsilon_n > 0$ consider the set

$$H_n(\varepsilon_n) := \{z \in S \,|\, h_n(z) \geq \varepsilon_n\}.$$

If

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

then $H_n(\varepsilon_n)$ is part of the set, where $|Q_{m_n}(z)| \leq e^{-n\varepsilon_n}$. By [4, Theorem 5.2.3] this latter set has logarithmic capacity $e^{-\varepsilon_n n/m_n}$, and hence (see [4, Theorem 5.3.5]) it has m_2-measure at most $\pi e^{-2\varepsilon_n n/m_n}$. Thus,

$$m_2(H_n(\varepsilon_n)) \leq \pi e^{-2\varepsilon_n n/m_n}.$$

Setting here $\varepsilon_n = \sqrt{m_n/n} \to 0$, we obtain

$$m_2(H_n(\varepsilon_n)) \leq \pi e^{-2\sqrt{n/m_n}},$$

hence there is a subsequence $\mathcal{N}_3 \subset \mathcal{N}_2$ such that

$$\sum_{n \in \mathcal{N}_3} m_2(H_n(\varepsilon_n)) < \infty.$$

Therefore, by the Borel-Cantelli lemma, m_2-almost all points $z \in H$ are contained in only finitely many of the sets $H_n(\varepsilon_n)$, $n \in \mathcal{N}_3$, and in all those points (6) is true.

■

After these preparations let $\nu_n = \nu_n^1 + \nu_n^2$, where ν_n^2 is the restriction of ν_n to the exterior of $\overline{B_{1/2}}$ (and hence ν_n^1 is the restriction of ν_n to $\overline{B_{1/2}}$). Let ν_n^3 be the balayage of ν_n^2 out of $\mathbf{C} \setminus \overline{B_{1/2}}$. Then, as before, ν_n^3 is a measure on $\partial B_{1/2}$ such that it has the same total mass as ν_n^2, and with some constant d_n we have

$$U^{\nu_n^2}(z) = U^{\nu_n^3}(z) + d_n, \qquad z \in \overline{B_{1/2}}.$$

Note however, that now we do not know if the total mass of ν_n^3/n tends to 0, all we know is that this measure has total mass at most 1 and it is supported on the circle $|z| = 1/2$. Set $\tilde{\nu}_n = \nu_n^1 + \nu_n^3$, for which

$$\frac{1}{n}U^{\nu_n}(z) - \frac{d_n}{n} = \frac{1}{n}U^{\tilde{\nu}_n}(z), \qquad z \in B_{1/2}. \tag{7}$$

Here $\tilde{\nu}_n$ have support in $\overline{B_{1/2}}$, and we may select a subsequence $\mathcal{N}_4 \subset \mathcal{N}_3$ such that along $\mathcal{N}_4$ the measures $\tilde{\nu}_n/n$ converge in the weak* topology to a measure $\tilde{\nu}$ supported on $\overline{B_{1/2}}$. Note that $\tilde{\nu}_n$ agrees with ν_n inside $B_{1/2}$ and ν_n/n was convergent along $\mathcal{N}_1$ to ν, so we get that ν and $\tilde{\nu}$ coincide inside $B_{1/2}$.

Now we invoke the lower envelope theorem (see [5, Theorem I.6.9]), according to which for all $z \in \mathbf{C}$, with the exception of a set of capacity 0, we have

$$\liminf_{n\to\infty,\ n\in\mathcal{N}_4} \frac{1}{n}U^{\tilde{\nu}_n}(z) = U^{\tilde{\nu}}(z). \tag{8}$$

In view of (4) and (5) there is a $z_0 \in S$ for which we have

$$\lim_{n\to\infty,\ n\in\mathcal{N}_2} \left(\frac{1}{n}U^{\mu_n}(z_0) - \frac{1}{n}U^{\nu_n}(z_0)\right) = 0, \tag{9}$$

$$\lim_{n\to\infty,\ n\in\mathcal{N}_3} \frac{1}{n}\left(U^{\mu_n}(z_0) - a_n\right) = U^{\mu}(z_0) \tag{10}$$

and (see (7) and (8))

$$\liminf_{n\to\infty,\ n\in\mathcal{N}_4} \left(\frac{1}{n}U^{\nu_n}(z_0) - \frac{d_n}{n}\right) = U^{\tilde{\nu}}(z_0),$$

where the right hand side is finite, i.e. along some subsequence $\mathcal{N}_5 \subset \mathcal{N}_4$

$$\lim_{n\to\infty,\ n\in\mathcal{N}_5} \left(\frac{1}{n}U^{\nu_n}(z_0) - \frac{d_n}{n}\right) = U^{\tilde{\nu}}(z_0). \tag{11}$$

Thus, along $\mathcal{N}_5$

$$\left(\frac{1}{n}U^{\mu_n}(z_0) - a_n\right) - \left(\frac{1}{n}U^{\nu_n}(z_0) - \frac{d_n}{n}\right) + a_n - \frac{d_n}{n} \to 0$$

(see (9)), and since the two expressions in the brackets also converge by (10) and (11) to a finite value, we obtain that $\{a_n - \frac{d_n}{n}\}$ converges (as $n \to \infty$, $n \in \mathcal{N}_5$), say it converges to the finite number b. Now, it follows from (4) and (7) that for m_2-almost all $z \in S$ we have

$$\left(\frac{1}{n}U^{\mu_n}(z) - a_n\right) - \frac{1}{n}U^{\tilde{\nu}_n}(z) + a_n - \frac{d_n}{n} \to 0,$$

along $\mathcal{N}_5$, and on invoking (5) we obtain that for almost all $z \in S$

$$\frac{1}{n}U^{\tilde{\nu}_n}(z) \to U^{\mu}(z) + b, \qquad \text{as } n \to \infty,\ n \in \mathcal{N}_5.$$

As a consequence, then

$$\liminf_{n\to\infty,\ n\in\mathcal{N}_5} \frac{1}{n}U^{\tilde{\nu}_n}(z) = U^{\mu}(z) + b$$

is also true on S m_2-almost everywhere. But, by the lower envelope theorem ([5, Theorem I.6.9]), the left hand side agrees with $U^{\tilde{\nu}}(z)$ everywhere except for a set of capacity 0 (in particular, m_2-almost everywhere), hence we finally obtain the equality

$$U^{\tilde{\nu}}(z) = U^{\mu}(z) + b \tag{12}$$

m_2-almost everywhere on S.

On taking the average of both sides in (12) over some small disk $B_r(z)$ about a fixed point $z \in S$, and letting r tend to 0 we obtain (12) everywhere on S, since, as $r \to 0$, we have, by the superharmonicity of logarithmic potentials,

$$\frac{1}{\pi r^2}\int_{B_r(z)} U^{\rho}(t)dt \to U^{\rho}(z)$$

for any measure ρ with compact support (cf. [4, Theorem 2.7.2] and its proof). Thus, (12) is true everywhere on S. In particular, since U^{μ} is harmonic in S, the same must be true of $U^{\tilde{\nu}}$, which implies that $\tilde{\nu}$ has no mass in S (see e.g. [4, Corollary 3.7.5]).

Let now γ be a C^2 Jordan curve in S that circles K once, and let ds be the arc measure on γ. We have just seen that all the mass of ν inside γ lies on K. If $\partial/\partial\mathbf{n}$ denotes normal derivative on γ in the direction of the inner normal, then, by Gauss' theorem (see [5, Theorem II.1.1]), the total mass of μ inside γ is

$$\mu(K) = \frac{1}{2\pi}\int_{\gamma} \frac{\partial U^{\mu}}{\partial\mathbf{n}} ds,$$

and the total mass of $\tilde{\nu}$ inside γ is

$$\tilde{\nu}(K) = \frac{1}{2\pi} \int_\gamma \frac{\partial U^{\tilde{\nu}}}{\partial \mathbf{n}} ds.$$

Since, by (12), here the right-hand sides are the same, we obtain

$$\tilde{\nu}(K) = \mu(K) = 1$$

which contradicts what we started with, i.e. with $\nu(K) \le \alpha < 1$, because $\tilde{\nu}(K) = \nu(K)$ (recall that ν and $\tilde{\nu}$ coincide inside $B_{1/2}$).

■

3 The Malamud-Pereira theorem

In 2003 an extension of the Gauss-Lucas theorem was found independently by S. M. Malamud [2] and R. Pereira [3]. To formulate their theorem let us recall that an $(n-1) \times n$ size $\mathcal{A} = (a_{ij})$ matrix is doubly stochastic if

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

Let p_n be a polynomial of degree n, let $z_1, \ldots, z_n$ be its zeros and let $\xi_1, \ldots, \xi_{n-1}$ the zeros of p_n'. Set

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

With these the Malamud-Pereira theorem states that there is a doubly stochastic

matrix $\mathcal{A}$ such that $\Xi = \mathcal{A}\mathbf{Z}$. An immediate consequence is that if $\varphi : \mathbf{C} \to \mathbf{R}_+$ is convex (in the classical sense that $\varphi(\alpha z + (1-\alpha)w) \le \alpha\varphi(z) + (1-\alpha)\varphi(w)$ for all z, w and $0 < \alpha < 1$), then

$$\frac{1}{n-1} \sum_{j=1}^{n-1} \varphi(\xi_j) \le \frac{1}{n} \sum_{k=1}^{n} \varphi(z_k). \tag{13}$$

Now we show that this implies Theorem 1 provided we know that all zeros of all p_n lie in a fixed compact set, say in the disk B_R. Indeed, consider a line L disjoint from K. It determines two half-planes, and let H_L be the half-plane which is disjoint from K. The claim in the theorem is easily seen to be equivalent to saying that there are $o(n)$ zeros of p_n' in every such H_L. To show that last claim, by the Gauss-Lucas theorem we may assume that L intersects B_R. We may also assume (apply rotation and translation) that L is the imaginary axis, and K lies to the left of the line $\Re z = -a$ with some $a > 0$. Consider the function $\varphi(z) = \max(0, \Re(z+a))$. This is convex, so we may apply (13). Since $\varphi(z) = 0$ on K, and $\varphi(z_k) \le 2R$ for all k (we wrote here $2R$ instead of R to allow for the just made translation and rotation), the right-hand side in (13) is at most $2Rm_n/n$, where m_n is the number of zeros of p_n lying outside K. Hence, by assumption, the right-hand side tends to 0, and therefore so does the left-hand side. However, on the left of (13) we have $\varphi(\xi_j) \ge a$ for every ξ_j lying in the right-half plane, which is H_L, and we obtain that there can be only $o(n)$ such ξ_j there.

Despite this simple proof, the Malamud-Pereira theorem does not seem to imply Theorem 1 in its full generality.

References

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

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

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

[4] T. Ransford, *Potential theory in the complex plane*, Cambridge University Press, Cambridge, 1995.

[5] E. B. Saff and V. Totik, *Logarithmic potentials with external fields*, Grundlehren der mathematischen Wissenschaften, **316**, Springer Verlag, Berlin, 1997.

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

and

Department of Mathematics and Statistics
University of South Florida
4202 E. Fowler Ave, CMC342
Tampa, FL 33620-5700, USA

totik@mail.usf.edu