A quantitative Gauss-Lucas theorem

Abstract

We prove in this note, a quantitative version of the classical Gauss-Lucas theorem: We show that for any degree n polynomial p and any $c \geq \frac{1}{2}$, the area of the convex hull of the roots of $p^{(cn)}$ is at most $4(c-c^2)$ that of the area of the convex hull of the roots of p. The proof uses the celebrated barrier method of Batson-Marcus-Spielman-Srivastava and a majorization theorem due independently to Pereira and Malamud.

1 Introduction

The fundamental Gauss-Lucas theorem[10][2.1] says that the critical points of a univariate polynomial lie in the convex hull of the polynomial's roots. Given a polynomial p and a positive integer k, we let $p^{(k)}$ denote the k'th derivative of p. We also use the notation $\sigma(p)$ to denote the roots of p and $\mathcal{K}(p)$ to denote the convex hull of the roots of p. Letting n be the degree of p, we have a nested collection of convex sets,

$$\mathcal{K}(p) \supset \mathcal{K}(p') \supset \mathcal{K}(p^{(2)}) \supset \cdots \supset \mathcal{K}(p^{(n-1)}).$$

It is easy to see that if let α be the average of the elements in $\sigma(p)$, the average of the elements in $\sigma(p^{(k)})$ equals α as well, for every $1 \leq k \leq n-1$. In particular, the convex sets $\mathcal{K}(p^{(k)})$ shrink to the one element set $\mathcal{K}(p^{(n-1)}) = \{\alpha\}$. It is natural to ask how quickly the sizes of these sets can shrink, something that we could not find a reference to in the literature. The main result in this paper is the following universal estimate, where given a set A in the plane, $|A|$ refers to the area of A.

Theorem 1 *Let p be a degree n polynomial. Then, for any $c \geq \dfrac{1}{2}$, we have that,*

$$|\mathcal{K}(p^{(\lceil cn \rceil)})| \leq 4(c-c^2)\, |\mathcal{K}(p)|.$$

Note that this estimate $4(c-c^2)$ is independent of the polynomial or even the degree n. These estimates are certainly not sharp but we suspect that the $O(1-c)$ dependance is. Also, by looking at the polynomial $p(z) = (z^3-1)^m$, one sees that one needs to take the derivative at least $\frac{n}{3} - 2$ times where $n = 3m$ is the degree of p, in order to get a shrinking of the areas of the convex hulls of higher derivatives. The theorem (1) as stated above cannot, by this simple observation, hold for $c \leq \frac{1}{3}$. It is conceivable that estimates could be got for c in the range $[\frac{1}{3}, \frac{1}{2}]$, but we do not do this in this paper.

We deduce this theorem from an analogous result for real rooted polynomials, that is of independent interest as it has applications to the restricted invertibility principle of Bourgain and Tzafriri[3], see [11] for a discussion. For real rooted polynomials, $\mathcal{K}(p)$ will refer to the difference of the largest and smallest roots of p, a quantity that is often called the *span* of the polynomial p.

Theorem 2 *Let p be a real rooted degree n polynomial. Then, for any $c \geq \frac{1}{2}$, we have that,*

$$|\mathcal{K}(p^{(\lceil cn \rceil)})| \leq 2\sqrt{c-c^2}\,|\mathcal{K}(p)|.$$

The estimates in the real-rooted case are tight, something that follows from a simple calculation. The proof of this theorem uses the barrier method introduced by Batson, Spielman and Srivastava [1] in their work on spectral sparsification and further developed by Marcus, Spielman and Srivastava [6] en route to their breakthrough results on Ramanujan graphs [7] and the solution of the Kadison-Singer problem [8].

We will then translate this result to the complex rooted case using the notion of majorization between real sequences by applying results of Pereira [9] and Malamud [5]. This will allow us to prove estimates on root shrinking in each direction. Deducing estimates on the shrinking of the areas of the convex hulls will then be a simple corollary.

2 Majorization relations for polynomial roots

We start off with a result proved by Pereira in 2005 [9] and conjectured by Katsoprinakis in the 1980's [4]. The result also appears in the contemporaneous work of Malamud [5] on closely related problems. Recall that a real sequence $\overline{\mu}$ is majorized by a real sequence $\overline{\lambda}$(of the same size), which we will denote $\overline{\mu} \prec \overline{\lambda}$ if there is a doubly stochastic map D such that $D\overline{\lambda} = \overline{\mu}$. Here, a doubly stoachastic map is a matrix of non-negative reals with all row and column sums 1. It is a classical fact that Majorization can also be expressed in terms of convex maps, in the following way, see [9][Prop. 4.2],

Theorem 3 *Let $\overline{\mu} = (\mu_1, \cdots, \mu_n)$ and $\overline{\lambda} = (\lambda_1, \cdots, \lambda_n)$ be two real sequences. Then, the following are equivalent,*

1. $\overline{\mu} \prec \overline{\lambda}$
2. *For every convex function f defined on an interval containing both $\overline{\lambda}$ and $\overline{\mu}$, we have that,*

$$\sum f(\mu_i) \leq \sum f(\lambda_i).$$

Given a polynomial p with roots $(\lambda_1, \cdots, \lambda_n)$, we will use the notation $R(p)$ to denote the monic polynomial whose roots are $(\text{Re } \lambda_1, \cdots, \text{Re } \lambda_n)$. The following was conjectured by Katsoprinakis[4] and proved 20 years later by Pereira [9][Theorem 4.6] and independently by Malamud [5],

Theorem 4 (Pereira) *Given a polynomial p, we have,*

$$\sigma\left(R(p')\right) \prec \sigma\left(R(p)'\right).$$

Given a real rooted polynomial p, recall that the potential function Φ_p on $[\lambda_{max}(p), \infty]$ was defined thus,

$$\Phi_p(b) := \sum \frac{1}{b - \lambda_k(p)}.$$

Recall too that we had called the inverse the soft max function of p, $\text{smax}_\varphi(p)$, that is,

$$\text{smax}_\varphi(p) := \Phi_p^{-1}(\varphi).$$

The function $f(x) = \dfrac{1}{b - x}$ is convex on $(-\infty, b)$ and by combining theorem (4) and proposition (7), we have that,

Theorem 5 *Given a polynomial p, we have, for any $b \geq \max \text{roots}\,(\tilde{p}')$,*

$$\Phi_{R(p')}(b) \leq \Phi_{R(p)'}(b).$$

And since the potential functions are monotone decreasing, we have that,

$$\text{smax}_\varphi(R(p')) \leq \text{smax}_\varphi\left(R(p)'\right).$$

We would like to point another interesting relation between roots of real parts of polynomials and their derivatives. It will be convenient to use the notation Dp to represent the derivative of p. The theorems of Pereira and Malamud show that,

$$\sigma(RD(p)) \prec \sigma(DR(p)). \tag{1}$$

We now show that there is an interesting extension for higher derivatives.

Theorem 6 *Let p be a degree n polynomial and let $k \leq n$. We then have a chain of majorization relations (between real sequences of size $n-k$),*

$$\sigma\left(RD^{(k)}(p)\right) \prec \sigma\left(DRD^{(k-1)}(p)\right) \prec \cdots \prec \sigma\left(D^{(k-1)}R(p)\right) \prec \sigma\left(D^{(k)}R(p)\right).$$

Proof. Applying (1) to the polynomial $D^{(k-1)}(p)$ yields that,

$$\sigma\left(RDD^{(k-1)}(p)\right) = \sigma\left(RD^{(k)}(p)\right) \prec \sigma\left(DRD^{(k-1)}(p)\right).$$

Borcea and Branden in [2][Theorem 1] showed(this is a very special case of their theorem) that if q and p are real rooted polynomials, then $\sigma(q) \prec \sigma(p)$ implies that $\sigma(Dq) \prec \sigma(Dp)$. Applying this to the polynomials $RD^{(k-1)}(p)$ and $DRD^{(k-2)}(p)$, we see, using (1) again that,

$$\sigma\left(DRD^{(k-1)}(p)\right) \prec \sigma\left(DDRD^{(k-2)}(p)\right) = \sigma\left(D^{(2)}RD^{(k-2)}(p)\right).$$

Iterating this argument establishes the theorem. ∎

3 The barrier method and the shrinking of root sets

Given a real rooted polynomial p, define the quantities,

$$\Phi_p(b) = \frac{p'}{p}(b), \; b > \lambda_{max}(p), \qquad \operatorname{smax}_\varphi(p) = \Phi^{-1}(\varphi), \; \varphi > 0.$$

The function $\operatorname{smax}_\varphi(p)$ which maps $(0, \infty) \to (\lambda_{max}(p), \infty)$ is a useful proxy for the max root of a polynomial and a discussion of its utility can be found in [6]. Two simple properties, see [6] or [11], that will be relevant are,

1. $\operatorname{smax}_\varphi(p)$ is positive, increasing and concave.
2. For every $\varphi < \infty$, we have that $\operatorname{smax}_\varphi(p) > \lambda_{max}$.

Another basic property of the above quantity is the following fact, which was announced by Adam Marcus in 2014. A proof can be found in the paper [11][Prop 3.1].

Proposition 7 (Marcus, 2014) *Let p be a real rooted polynomial and $\varphi \in (0, \infty]$. Then,*

$$\operatorname{smax}_\varphi(p') \leq \operatorname{smax}_\varphi(p) - \frac{1}{\varphi}.$$

Together with theorem (5), we conclude that,

$$\operatorname{smax}_\varphi(R(p')) \leq \operatorname{smax}_\varphi(R(p)') \leq \operatorname{smax}_\varphi(R(p)) - \frac{1}{\varphi}.$$

Iterating this, we have that,

$$\operatorname{smax}_\varphi(R(p^{(cn)})) \leq \operatorname{smax}_\varphi(R(p)) - \frac{cn}{\varphi}.$$

We now perform some routine optimization,

Lemma 8 *Let p be a polynomial of degree n with roots lying in $B(0, 1)$ and with the average of the roots β. Then, letting $\alpha = \operatorname{Re} \beta$, for any $c \geq \frac{1+\alpha}{2}$,*

$$\lambda_{max} R(p^{(cn)}) \leq \left(\sqrt{(1-\alpha)(1-c)} + \sqrt{(1+\alpha)\, c}\right)^2 - 1.$$

Proof. We see that,

$$\lambda_{max} R(p^{(cn)}) \leq \inf_{\varphi \geq 0} \operatorname{smax}_\varphi(R(p^{(cn)})) \leq \operatorname{smax}_\varphi(R(p)) - \frac{cn}{\varphi} \leq \inf_{b>1} b - \frac{cn}{\Phi_p(b)},$$

with the last inequality following from the condition that the roots of $R(p)$ are at most 1. We also have that,

$$\Phi_p(b) = \sum_{\lambda \in \sigma(p)} \frac{1}{b - \lambda},$$

and it is easy to see, noting that the roots of $R(p)$ lie in $(-1,1)$ that the last expression is at most $\dfrac{n(1+\alpha)}{2(b-1)}+\dfrac{n(1-\alpha)}{2(b+1)}=\dfrac{n(b+\alpha)}{b^2-1}$. A simple calculation shows that,

$$\inf_{b>1} b-\frac{c(b^2-1)}{b+\alpha}=\begin{cases}\left(\sqrt{(1-\alpha)(1-c)}+\sqrt{(1+\alpha)\,c}\right)^2-1, & c\geq\dfrac{1+\alpha}{2},\\ 1, & c\leq\dfrac{1+\alpha}{2}.\end{cases}$$

The lemma follows. ∎

We will deduce another simple lemma from this,

Lemma 9 *Let p be a polynomial of degree n. Then, for any $c\geq\frac{1}{2}$, we have, letting $|\mathcal{K}(R(p))|=\lambda_{max}(R(p))-\lambda_{min}(R(p))$, that,*

$$|\sigma(R(p^{(cn)}))|\leq 2\sqrt{c-c^2}\,|\sigma(R(p))|.$$

Proof. It is easy to see that shifting and scaling the roots of the polynomial p does not affect the ratio $\dfrac{|\sigma(R(p^{(cn)}))|}{|\sigma(q)|}$ and we may therefore assume that the polynomial p has roots in $B(0,1)$. Let α be the average of the roots of $R(p)$. Applying lemma(8), we see that,

$$\lambda_{max}(R(p^{(cn)})))\leq\left(\sqrt{(1-\alpha)(1-c)}+\sqrt{(1+\alpha)\,c}\right)^2-1,\quad\text{if } c\geq\frac{1+\alpha}{2}.$$

Working with the polynomial $q(x)=p(-x)$, we have that the average of the roots of $R(q)$ is $-\alpha$, we see that,

$$\lambda_{min}(R(p^{(cn)}))\geq 1-\left(\sqrt{(1+\alpha)(1-c)}+\sqrt{(1-\alpha)\,c}\right)^2,\quad\text{if } c\geq\frac{1-\alpha}{2}.$$

Without loss of generality, we may assume that $\alpha\leq 0$ (else we work with r instead. We therefore have that,

$$|\mathcal{K}(R(p^{(cn)}))|\leq\begin{cases}4\sqrt{c(1-c)(1-\alpha^2)}, & c\geq\dfrac{1-\alpha}{2}\\ \left(\sqrt{(1-\alpha)(1-c)}+\sqrt{(1+\alpha)\,c}\right)^2, & \dfrac{1+\alpha}{2}\leq c\leq\dfrac{1-\alpha}{2}\end{cases}$$

In the case when $c\geq\dfrac{1-\alpha}{2}$, we note that the expression $4\sqrt{c(1-c)(1-\alpha^2)}$ is maximized when $\alpha=0$ where it equals $4\sqrt{c(1-c)}$.

For the case when $\dfrac{1+\alpha}{2}\leq c\leq\dfrac{1-\alpha}{2}$, but still, $c\geq\dfrac{1}{2}$: We see that for fixed c, the expression,

$$\left(\sqrt{(1-\alpha)(1-c)}+\sqrt{(1+\alpha)\,c}\right)^2, \tag{2}$$

as a function of α increases from -1 to $2c-1$ and then decreases from $2c-1$ to 1. We have the condition $\dfrac{1+\alpha}{2}\leq c\leq\dfrac{1-\alpha}{2}$ which gives us that $\alpha\leq\min\{2c-1,1-2c\}$.

Together with the condition $c \geq \frac{1}{2}$, this reduces to the condition $\alpha \leq 1-2c$. The expression (2) subject to this constraint on α thus has a maximum value at $\alpha = 1-2c$, where it equals $8c(1-c)$. It is easy to see that this is smaller than $4\sqrt{c(1-c)}$ for every $c \in [0,1]$. And finally, using the fact $R(p)$ has roots in $(-1,1)$,

$$\frac{|\mathcal{K}(R(p^{(cn)}))|}{|\mathcal{K}(R(p))|} \leq \frac{4\sqrt{c(1-c)}}{2} = 2\sqrt{c(1-c)}.$$

■

We now deduce our main result, a quantitative Gauss-Lucas theorem,

Theorem 10 *Let p be a polynomial of degree n. Then, for any $c \geq \frac{1}{2}$, we have that,*

$$|\mathcal{K}(p^{(cn)})| \leq 4(c-c^2)\, |\sigma(p)|.$$

Proof. Lemma (9) says that the ratio between the sizes of the projections of $\sigma(p^{(cn)})$ and $\sigma(p)$ onto the real axis is at most $2\sqrt{c-c^2}$. There is nothing special about the real axis; Working with $q(z) = p(e^{-i\theta}z)$, we see that the ratios of the projections onto the line $\mathrm{Arg}(z) = \theta$ are again bounded by $2\sqrt{c-c^2}$. We therefore have two polygons with the properties,

1. The ratios of their shadows in every direction are at most $2\sqrt{c-c^2}$.
2. They have the same centroid(since the roots of a polynomial and its critical points have the same average).

Writing out the areas in polar coordinates shows that the ratio of the areas is at most $4(c-c^2)$. ■

Let us mention another result along these lines.

Theorem 11 *Let p be a degree n polynomial with roots in $B(0,1)$ and with average of its roots 0. Then, for any $c \geq \frac{1}{2}$,*

$$\sigma(p^{(cn)}) \subset B(0, 2\sqrt{c-c^2}).$$

Proof. The real rooted polynomial $R(p)$ has roots in $(-1,1)$ and the average of its roots is 0. Lemma(8) then implies that,

$$\sigma\left(R(p^{(cn)})\right) \subset (-2\sqrt{c-c^2}, 2\sqrt{c-c^2}).$$

And clearly, the same holds for any other lie that we project the roots to. The theorem follows. ■

4 Tightness of bounds

Lemma (9) implies that when the polynomial p is real rooted, we have, letting $|\mathcal{K}(p)|$ be the size of the smallest interval containing $\sigma(p)$, that,

$$|\mathcal{K}\left(p^{(cn)}\right)| \leq 2\sqrt{c-c^2}|\mathcal{K}(p)|.$$

This is sharp: The polynomial $(z^2-1)^m$ shows that one needs to take the derivative at least $\frac{n}{2}$ times where $n=2m$ to have all the roots migrate inward from the end points. Further, a simple calculation involving comparing coefficients shows that,

$$\sum_{\lambda\in\sigma(p^{(cn)})}\lambda^2=\frac{(n-cn)(n-cn-1)}{n(n-1)}=n(1-c)^2-\frac{nc}{n-1}.$$

This implies that there is at least one root of modulus at least $\sqrt{1-c}-O(\frac{1}{n})$ and since the roots of $p^{(cn)}$ are symmetric about 0, the smallest interval containing all the roots of $p^{(cn)}$ contains $[-\sqrt{1-c}+O(\frac{1}{n}),\sqrt{1-c)}-O(\frac{1}{n})]$. We conclude that in the class of real rooted polynomials, which we denote $\mathcal{Q}$ and for any $c\geq\frac{1}{2}$,

$$\inf_{p\in\mathcal{Q}}\frac{|\mathcal{K}\left(p^{(cn)}\right)|}{|\mathcal{K}(p)|}\geq\sqrt{1-c}.$$

This shows that the upper bound from theorem (9) is optimal upto a constant. For the complex rooted case, we make an analogous calculation with the polynomial $(z^3-1)^n$. We have,

$$p(z)=z^{3n}-nz^{3n-3}+\cdots,$$

and

$$p^{(3cn)}(z)=\binom{3n}{3n(1-c)}z^{3(1-c)n}-n\binom{3n-3}{3n(1-c)-3}z^{3n-3}+\cdots.$$

The polynomial $p^{(3cn)}$ has roots of the form $\{\lambda_i,\lambda_i\omega,\lambda_i\omega^2::1\leq i\leq(1-c)n\}$ where the λ_i are non-negative reals and we have that,

$$p^{(3cn)}=\prod_{i=1}^{(1-c)n}(z^3-\lambda_i^3).$$

Comparing coefficients, we see that,

$$\sum_{i=1}^{cn}\lambda_i^3=n\binom{3n-3}{3n(1-c)-3}/\binom{3n}{3n(1-c)}=n(1-c)^3+O\left(\frac{1}{n}\right).$$

The largest of the λ_i, which we may assume is λ_1, is therefore at least $(1-c)+O(\frac{1}{n})$. The convex hull of the roots of $p^{(3cn)}$ is the equilateral triangle with vertices $\{\lambda_1,\lambda_1\omega,\lambda_1\omega^2\}$ and we see that,

$$\frac{|\mathcal{K}(p^{(cn)})|}{|\mathcal{K}(p)|}\geq(1-c)^{4/3}+O\left(\frac{1}{n}\right).$$

We conclude that, letting $\mathcal{P}$ be the class of all polynomials and working with areas of the convex hulls,

$$\inf_{p\in\mathcal{P}}\frac{|\mathcal{K}\left(p^{(cn)}\right)|}{|\mathcal{K}(p)|}\geq(1-c)^{4/3}.$$

I suspect this can be improved to $O(1-c)$ to match the upper bound.

References

[1] Joshua Batson, Daniel A. Spielman, and Nikhil Srivastava, *Twice-Ramanujan sparsifiers*, SIAM J. Comput. **41** (2012), no. 6, 1704–1721. MR 3029269

[2] Julius Borcea and Petter Brändén, *Hyperbolicity preservers and majorization*, C. R. Math. Acad. Sci. Paris **348** (2010), no. 15-16, 843–846. MR 2677977

[3] J. Bourgain and L. Tzafriri, *On a problem of Kadison and Singer*, J. Reine Angew. Math. **420** (1991), 1–43. MR 1124564

[4] Emmanuel S. Katsoprinakis, *On the complex Rolle set of a polynomial*, Computational methods and function theory 1997 (Nicosia), Ser. Approx. Decompos., vol. 11, World Sci. Publ., River Edge, NJ, 1999, pp. 315–326. MR 1700357

[5] S. M. Malamud, *Inverse spectral problem for normal matrices and the Gauss-Lucas theorem*, Trans. Amer. Math. Soc. **357** (2005), no. 10, 4043–4064. MR 2159699

[6] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava, *Ramanujan graphs and the solution of the Kadison-Singer problem*, Proc. ICM **3** (2014), 375–386.

[7] ______, *Interlacing families I: Bipartite Ramanujan graphs of all degrees*, Ann. of Math. (2) **182** (2015), no. 1, 307–325. MR 3374962

[8] ______, *Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem*, Ann. of Math. (2) **182** (2015), no. 1, 327–350. MR 3374963

[9] Rajesh Pereira, *Differentiators and the geometry of polynomials*, J. Math. Anal. Appl. **285** (2003), no. 1, 336–348. MR 2000158

[10] Q. I. Rahman and G. Schmeisser, *Analytic theory of polynomials*, London Mathematical Society Monographs. New Series, vol. 26, The Clarendon Press, Oxford University Press, Oxford, 2002. MR 1954841

[11] Mohan Ravichandran, *Principal submatrices and restricted invertibility*, https://arxiv.org/1609.04187v1.