

A CONTRACTION OF THE LUCAS POLYGON

BRANKO ČURĀUS AND VANIA MASCIONI

(Communicated by N. Tomczak-Jaegermann)

ABSTRACT. The celebrated Gauss-Lucas theorem states that all the roots of the derivative of a complex non-constant polynomial p lie in the convex hull of the roots of p , called the Lucas polygon of p . We improve the Gauss-Lucas theorem by proving that all the nontrivial roots of p' lie in a smaller convex polygon which is obtained by a strict contraction of the Lucas polygon of p .

Based on a simple proof of the classical Gauss-Lucas theorem from [5, Theorem 4.4.1] and an inequality from [2], we give an improvement of this theorem by showing that all the nontrivial roots of the derivative of a polynomial lie in a smaller convex polygon than predicted by the Gauss-Lucas theorem. In fact, our result is closely related to the following consideration of J. L. Walsh.

In [12, §3.4] J. L. Walsh wrote: “A deleted neighborhood of an arbitrary zero of $p(z)$ can be assigned which is known to contain no critical point of $p(z)$. Since no critical point other than a multiple zero of $p(z)$ can lie on the boundary of the Lucas polygon (assumed non-degenerate), it follows that no critical point lies in a certain strip inside the polygon and bounded by a side of the polygon and by a line parallel to that side. We proceed to make this conclusion more precise under certain conditions: [...]”

However, Walsh’s results as described in [12] improve the Lucas polygon only along sides spanned by two roots of the same multiplicity and not containing any other root in the interior. Also, the formulas and calculations required in order to practically compute his improvement seem prohibitive.

A significant improvement of the Gauss-Lucas theorem was achieved by D. Dimitrov in [3]. The regions that are guaranteed to contain all the nontrivial critical points of a polynomial p in [3] are given as intersections of certain unions of circles. The main difference between the results in [3] and our result is that the regions in [3] are not easy to visualize, they are not convex and there is no guarantee that these regions are strictly contained in the Gauss-Lucas polygon. In fact, a calculation of the regions studied in [3, Corollary 1] shows that in none of the cases studied in Example 3 below are these regions strictly included in the corresponding contracted Gauss-Lucas polygons constructed in this note. Therefore, the best prediction for the location of the nontrivial critical points of a polynomial is obtained as the intersection of the regions from [3] and our contracted Gauss-Lucas polygon.

Received by the editors October 29, 2002 and, in revised form, February 12, 2003.

2000 *Mathematics Subject Classification.* Primary 30C15; Secondary 26C10.

Key words and phrases. Roots of polynomials, critical points of polynomials, Gauss-Lucas theorem.

What we set out to do in this paper (see Theorem 2) is to furnish a much easier procedure that allows us to give a simple specific description of “a certain strip” from Walsh’s quote for any polynomial p with at least two distinct roots, and this strip actually applies to the entire boundary of the Lucas polygon.

For a historic account on the Lucas theorem see [6, Section 2]. For applications of the Lucas theorem in different settings see [1, 4, 7, 8, 9, 10, 11, 13].

We consider complex polynomials. By \bar{z} we denote the complex conjugate of $z \in \mathbb{C}$ and $D(u, r)$ denotes a closed disk in \mathbb{C} with radius $r > 0$ and center $u \in \mathbb{C}$.

Theorem 1 (Gauss-Lucas theorem). *Let p be a non-constant polynomial with derivative p' . All the roots of p' lie in the convex hull of the roots of p .*

We include an explicit proof of this theorem because this allows us to introduce notation and equations useful in the sequel. We note that this unusually simple proof, while not common in recent books, is not new: see, for instance, [5, Theorem 4.4.1] or [13, Theorem 2.2.9].

Proof. Let ζ be a nontrivial root of p' ; that is, let $p'(\zeta) = 0$ and $p(\zeta) \neq 0$. Let w_1, \dots, w_k be all the (distinct) roots of p with multiplicities m_1, \dots, m_k , respectively. It is no restriction to assume that p is of the form

$$(1) \quad p(z) = \prod_{j=1}^k (z - w_j)^{m_j}.$$

Taking the derivative of (1) and dividing by $p(z) \neq 0$ we get

$$(2) \quad \frac{p'(z)}{p(z)} = \sum_{j=1}^k \frac{m_j}{z - w_j}.$$

Substituting ζ in (2) and using the fact that $\bar{\bar{0}} = 0$ we obtain

$$\begin{aligned} 0 &= \frac{p'(\zeta)}{p(\zeta)} = \sum_{j=1}^k \frac{m_j}{\zeta - w_j} = \sum_{j=1}^k \frac{m_j}{\bar{\zeta} - \bar{w}_j} \\ &= \sum_{j=1}^k \frac{m_j}{\bar{\zeta} - \bar{w}_j} \frac{|\zeta - w_j|^2}{|\zeta - w_j|^2} = \sum_{j=1}^k \frac{m_j}{|\zeta - w_j|^2} (\zeta - w_j). \end{aligned}$$

Thus

$$(3) \quad 0 = \sum_{j=1}^k \frac{m_j}{|\zeta - w_j|^2} (\zeta - w_j) = \sum_{j=1}^k c_j (\zeta - w_j),$$

where we put

$$c_j := \frac{m_j}{|\zeta - w_j|^2} > 0 \quad j = 1, \dots, k.$$

Solving (3) for ζ we get

$$(4) \quad \zeta = \frac{1}{\sum_{j=1}^k c_j} \left(\sum_{j=1}^k c_j w_j \right) = \sum_{j=1}^k d_j w_j,$$

where

$$(5) \quad d_j := \frac{c_j}{\sum_{\mu=1}^k c_\mu}, \quad j = 1, \dots, k.$$

Clearly $0 < d_j \leq 1$, $j = 1, \dots, k$, and $\sum_{j=1}^k d_j = 1$. Thus, (4) shows that ζ is in the convex hull of the roots of p and the theorem is proved. \square

Next we define the quantities that we will need to formulate and prove an improvement of the Gauss-Lucas theorem. Let p be a polynomial of degree n , $n \geq 2$, and assume that p has at least two distinct roots. Let $Z(p)$ denote the set of all roots of p . Define

$$\begin{aligned}\omega(p, w) &:= \min\{|w - v| : v \in Z(p), w \neq v\}, \quad w \in Z(p), \\ \omega(p) &:= \min\{|w - v| : w, v \in Z(p), w \neq v\}, \\ \Omega(p) &:= \max\{|w - v| : w, v \in Z(p), w \neq v\}, \\ \tau(p, w) &:= \min\{|w - v| : v \in Z(p'), v \neq w\}, \quad w \in Z(p), \\ \tau(p) &:= \min\{|w - v| : w \in Z(p), v \in Z(p'), v \neq w\}, \\ T(p) &:= \max\{|w - v| : w \in Z(p), v \in Z(p'), v \neq w\}.\end{aligned}$$

In [2, Theorem 4] we proved the following inequalities:

$$(6) \quad 0 < \frac{1}{n} \omega(p) \leq \tau(p) \leq \frac{1}{2 \sin(\pi/n)} \omega(p)$$

and

$$(7) \quad 0 < \frac{m_w}{n} \omega(p, w) \leq \tau(p, w) \leq \frac{1}{\sin(\pi/(n - m_w))} \omega(p, w),$$

where m_w is the multiplicity of $w \in Z(p)$.

We will also need the following elementary estimate:

$$(8) \quad T(p) \leq \sqrt{\Omega(p)^2 - \tau(p)^2}.$$

To see this, let $w \in Z(p)$ and $\zeta \in Z(p')$, $\zeta \neq w$, be such that $|w - \zeta| = T(p)$. If ℓ is the line passing through ζ and perpendicular to the segment $w\zeta$, then Lucas' theorem implies that some $v \in Z(p)$ exists in the half-plane bounded by ℓ and not containing w . Since v does not lie inside the disk $D(v, \tau(p))$, it follows that

$$\Omega^2 \geq |w - v|^2 \geq T(p)^2 + \tau(p)^2,$$

and so (8) follows. Using (8) and (6) now gives

$$(9) \quad T(p) \leq \sqrt{\Omega(p)^2 - \omega(p)^2/n^2}.$$

Theorem 2. *Let p be a non-constant polynomial of degree n with derivative p' . Let w_1, \dots, w_k be all the (distinct) roots of p with multiplicities m_1, \dots, m_k , respectively. Put $\delta = \sum_{j=1}^k \delta_j$, where*

$$\delta_j := \frac{m_j}{m_j + (n^2 \Omega(p)^2 - \omega(p)^2) \sum_{i \neq j} \frac{1}{m_i \omega(p, w_i)^2}},$$

and define

$$w_b := \frac{1}{\delta} \sum_{j=1}^k \delta_j w_j.$$

Then $0 < \delta < 1$ and all the nontrivial roots of p' (that is, the roots of p' that are not the roots of p) lie in the "Swiss cheese"-like region obtained by removing the interiors of all the disks $D(w_j, m_j \omega(p, w_j)/n)$, $j = 1, \dots, k$, from the convex

polygon which is the contraction of the Lucas polygon of p centered at w and with the contraction coefficient $1 - \delta$.

Proof. Since the contraction centered at w_b with the contraction coefficient $1 - \delta$ is an affine bijective transformation of the complex plane, it maps the Lucas polygon of p onto the convex hull of the points

$$v_j := \delta w_b + (1 - \delta)w_j, \quad j = 1, \dots, k.$$

To prove that this smaller polygon contains all the roots of p' , we give detailed estimates for the coefficients c_j and d_j , $j = 1, \dots, k$, from the proof of the Gauss-Lucas theorem. Put $\omega(p) = \omega$ and $\Omega(p) = \Omega$. By the definitions and with ζ and d_j as in (4) and (5), we have (for every $j = 1, \dots, k$)

$$(10) \quad d_j = \frac{c_j}{\sum_{\mu=1}^k c_\mu} = \frac{m_j}{n} \left[\frac{n}{m_j + |\zeta - w_j|^2 \sum_{i \neq j} \frac{m_i}{|\zeta - w_i|^2}} \right].$$

Since by (9),

$$|\zeta - w_j| \leq T(p) \leq \sqrt{\Omega(p)^2 - \omega(p)^2/n^2}$$

and by (7),

$$|\zeta - w_i| \geq \tau(p, w_i) \geq \frac{m_i \omega(p, w_i)}{n},$$

the equality (10) implies

$$(11) \quad d_j \geq \frac{m_j}{m_j + (n^2 \Omega(p)^2 - \omega(p)^2) \sum_{i \neq j} \frac{1}{m_i \omega(p, w_i)^2}} =: \delta_j.$$

It follows from (11) and the definition of d_j in (5) (or (10)) that

$$(12) \quad 0 < \delta_j \leq d_j \leq 1, \quad j = 1, \dots, k,$$

and $\sum_{\mu=1}^k \delta_\mu = \delta < 1$. Also, for an arbitrary $j \in \{1, \dots, k\}$ we have

$$(13) \quad d_j = 1 - \sum_{\substack{\mu=1 \\ \mu \neq j}}^k d_\mu \leq 1 - \sum_{\substack{\mu=1 \\ \mu \neq j}}^k \delta_\mu = 1 - \delta + \delta_j < 1.$$

By (4), (12) and (13) all the nontrivial roots of p' lie in the convex region

$$R_w := \left\{ \sum_{\mu=1}^k t_\mu w_\mu : \delta_\mu \leq t_\mu \leq 1 - \delta + \delta_\mu, \sum_{\mu=1}^k t_\mu = 1 \right\}.$$

Denote by R_v the convex hull of the points

$$v_j := \delta w_b + (1 - \delta)w_j = \sum_{\substack{\mu=1 \\ \mu \neq j}}^k \delta_\mu w_\mu + (1 - \delta + \delta_j)w_j, \quad j = 1, \dots, k,$$

that is,

$$R_v := \left\{ \sum_{\mu=1}^k s_\mu v_\mu : 0 \leq s_\mu \leq 1, \sum_{\mu=1}^k s_\mu = 1 \right\}.$$

To prove that $R_w = R_v$ we will establish a one-to-one correspondence between the simplices

$$S_w = \left\{ [t_1 \ t_2 \ t_3 \ \dots \ t_{k-1} \ t_k]^T : \delta_\mu \leq t_\mu \leq 1 - \delta + \delta_\mu, \sum_{\mu=1}^k t_\mu = 1 \right\}$$

and

$$S_v = \left\{ [s_1 \ s_2 \ s_3 \ \cdots \ s_{k-1} \ s_k]^T : 0 \leq s_\mu \leq 1, \sum_{\mu=1}^k s_\mu = 1 \right\}$$

such that

$$\sum_{\mu=1}^k t_\mu w_\mu = \sum_{\mu=1}^k s_\mu v_\mu.$$

This will be accomplished by solving the system $\mathbf{A}\mathbf{s} = \mathbf{t}$ where

$$\mathbf{A} = \begin{bmatrix} 1 - \delta + \delta_1 & \delta_1 & \delta_1 & \cdots & \delta_1 & \delta_1 \\ \delta_2 & 1 - \delta + \delta_2 & \delta_2 & \cdots & \delta_2 & \delta_2 \\ \delta_3 & \delta_3 & 1 - \delta + \delta_3 & \cdots & \delta_3 & \delta_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \delta_{k-1} & \delta_{k-1} & \delta_{k-1} & \cdots & 1 - \delta + \delta_{k-1} & \delta_{k-1} \\ \delta_k & \delta_k & \delta_k & \cdots & \delta_k & 1 - \delta + \delta_k \end{bmatrix}$$

and $\mathbf{s} = [s_1 \ s_2 \ s_3 \ \cdots \ s_{k-1} \ s_k]^T$, $\mathbf{t} = [t_1 \ t_2 \ t_3 \ \cdots \ t_{k-1} \ t_k]^T$. The matrix \mathbf{A} is invertible with the inverse

$$\mathbf{A}^{-1} = \frac{1}{1 - \delta} \begin{bmatrix} 1 - \delta_1 & -\delta_1 & -\delta_1 & \cdots & -\delta_1 & -\delta_1 \\ -\delta_2 & 1 - \delta_2 & -\delta_2 & \cdots & -\delta_2 & -\delta_2 \\ -\delta_3 & -\delta_3 & 1 - \delta_3 & \cdots & -\delta_3 & -\delta_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\delta_{k-1} & -\delta_{k-1} & -\delta_{k-1} & \cdots & 1 - \delta_{k-1} & -\delta_{k-1} \\ -\delta_k & -\delta_k & -\delta_k & \cdots & -\delta_k & 1 - \delta_k \end{bmatrix}.$$

Therefore, the entries of $\mathbf{s} = \mathbf{A}^{-1}\mathbf{t}$ for $\mathbf{t} \in S_w$ are

$$(14) \quad s_j = \frac{t_j - \delta_j}{1 - \delta}, \quad j = 1, \dots, k.$$

The equalities in (14) imply that \mathbf{A}^{-1} maps the simplex S_w into the simplex S_v . Conversely, calculating the components of $\mathbf{t} = \mathbf{A}\mathbf{s}$ for $\mathbf{s} \in S_v$ we get

$$(15) \quad t_j = (1 - \delta)s_j + \delta_j, \quad j = 1, \dots, k,$$

and conclude that \mathbf{A} maps the simplex S_v into S_w . Thus \mathbf{A} is a bijection between S_v and S_w . This proves that all the nontrivial roots of p' lie in the convex hull R_v of the points v_1, \dots, v_k , which is the contraction of the Lucas polygon of p centered at w_b and with the contraction coefficient $1 - \delta$.

To see the part of the statement about “Swiss cheese”-like region, note that the definition of $\tau(p, w)$ implies that none of the disks $D(w, \tau(p, w))$, $w \in Z(p)$, contains nontrivial roots of p' . By the inequality (7) the same is true for the smaller disks $D(w, m_w \omega(p, w)/n)$. This completes the proof of the theorem. \square

Example 3. Let $u := \frac{1}{2} + i\frac{\sqrt{3}}{2}$. In this example and in Figures 1 through 6 we illustrate Theorem 2 by polynomials with roots of various multiplicities at the points $1, u, \bar{u}$ and 0 . Figures 1 through 6 show the original Lucas polygon, the shrunken Lucas polygon and the “Swiss” cheese-like region that contains the roots of the derivative as described in Theorem 2 of the polynomials given in the captions. The roots of the polynomials are marked by small black disks and the nontrivial roots of the derivatives are marked by small circles. More details for each figure are given below.

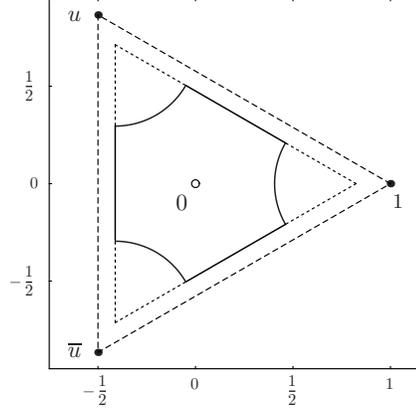
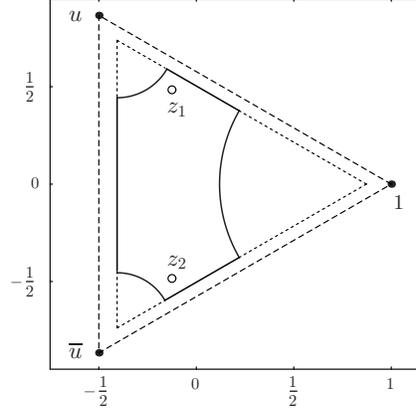
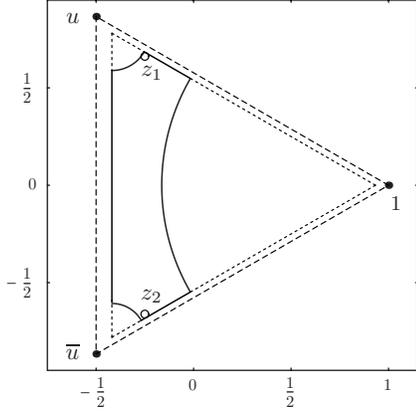
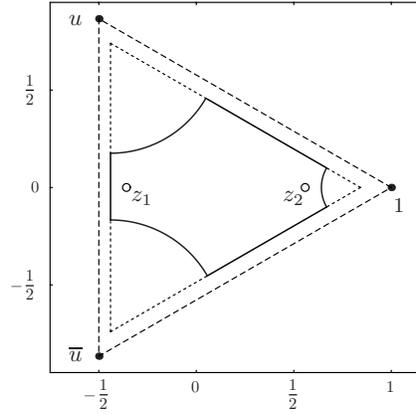
FIGURE 1. $1 - z^3$ FIGURE 2. $1 - z - z^3 + z^4$ FIGURE 3.
 $(z - 1)^4(z - u)(z - \bar{u})$ FIGURE 4.
 $1 + z + z^2 - z^3 - z^4 - z^5$

Figure 1. The polynomial is $1 - z^3$ with the simple roots $1, u, \bar{u}$ and 0 as a double root of the derivative. Easy computations give $\omega(p) = \omega(p, w_j) = \Omega(p) = \sqrt{3}$, and so $\delta_j = 1/17$ and $\delta = 3/17$. Since $w_b = 0$, by Theorem 2 the two nontrivial roots must lie inside the contraction of the Lucas polygon of p by the factor $14/17$ and the “Swiss-cheese” remark allows us to shave three small circular triangles from the corners of this smaller region.

Figure 2. The polynomial is $1 - z - z^3 + z^4$ with the double root 1 , and single roots u and \bar{u} . The derivative $-1 - 3z^2 + 4z^3$ has roots

$$z_1 = -\frac{1}{8} + i\frac{\sqrt{15}}{8}, \quad z_2 = -\frac{1}{8} - i\frac{\sqrt{15}}{8}, \quad z_3 = 1.$$

We calculate that $\delta = 111/752$, the contraction coefficient $1 - \delta = 641/752 \approx 0.85239$, and $w_b = 5/37$.

Figure 3. The polynomial is $1 - 3z + 3z^2 - 2z^3 + 3z^4 - 3z^5 + z^6$ with the quadruple root 1, and single roots u and \bar{u} . The derivative $3(-1+z)^3(1+z+2z^2)$ has nontrivial roots

$$z_1 = -\frac{1}{4} + i\frac{\sqrt{7}}{4}, \quad z_2 = -\frac{1}{4} - i\frac{\sqrt{7}}{4}.$$

We calculate that $\delta = 654/6623$, the contraction coefficient $1 - \delta = 5969/6623 \approx 0.90125$, and $w_b = 35/109$. The distance of the root z_1 to the closest edge of the contracted Lucas polygon is ≈ 0.01866 and its distance to the arc next to it is ≈ 0.03437 .

Figure 4. The polynomial is $1 + z + z^2 - z^3 - z^4 - z^5$ with the single root at 1 and double roots at u and \bar{u} . Its derivative $1 + 2z - 3z^2 - 4z^3 - 5z^4$ has roots

$$z_1 = \frac{1 - \sqrt{21}}{10}, \quad z_2 = \frac{1 + \sqrt{21}}{10}, \quad z_3 = u, \quad z_4 = \bar{u}.$$

We calculate that $\delta = 69/475$, the contraction coefficient $1 - \delta = 406/475 \approx 0.85474$, and $w_b = -2/23$.

Figure 5. The polynomial is

$$-1 + i\sqrt{3} - 2i\sqrt{3}(z - z^4) + (3 + i\sqrt{3})(z^2 - z^5) - (1 + i\sqrt{3})z^3 + 2z^6$$

with a triple root at 1, a double root at u and a single root at \bar{u} . The derivative is

$$(z - 1)^2(z - u)(3 - i\sqrt{3} + (3 + i\sqrt{3})z + 12z^2)$$

with the nontrivial roots

$$\begin{aligned} z_{1,2} &= -\frac{1}{8} - i\frac{1}{24}\left(\sqrt{3} \pm \sqrt{138 - 54i\sqrt{3}}\right) \\ &\approx -0.125 - 0.07217i \mp (0.15786 + 0.5143i). \end{aligned}$$

We calculate that $\delta = 62723/488881$, the contraction coefficient

$$1 - \delta = \frac{426158}{488881} \approx 0.8717,$$

and

$$w_b = 35(473 + 111i\sqrt{3})/125446 \approx 0.13197 + 0.05364i.$$

Figure 6. The polynomial is $z - z^4$ with the single roots at 0, 1, u and \bar{u} . The derivative is $1 - 4z^3$ with the roots

$$\frac{1}{\sqrt[3]{4}}, \quad \frac{1}{\sqrt[3]{4}}u, \quad \frac{1}{\sqrt[3]{4}}\bar{u}.$$

We calculate that $\delta = 2/71$, the contraction coefficient $1 - \delta = 69/71 \approx 0.97183$, and $w_b = 0$. Note that the region predicted by Theorem 2 to contain all the roots of the derivative consists of the intersection of the complement of the disk centered at 0, bounded by the thick circle and the region bounded by the thick line in Figure 6.

Example 4. Consider $p(z) = z^{n-1}(z - 1)$, and let $w_1 = 0, w_2 = 1$ (with multiplicities $m_1 = n - 1, m_2 = 1$). Easy computations give $\omega(p) = \omega(p, w_j) = \Omega(p) = 1$. So, in this case we have $w_b = 1/2$,

$$\delta_1 = \delta_2 = \frac{1}{n+2} \quad \text{and} \quad \delta = \frac{2}{n+2}.$$

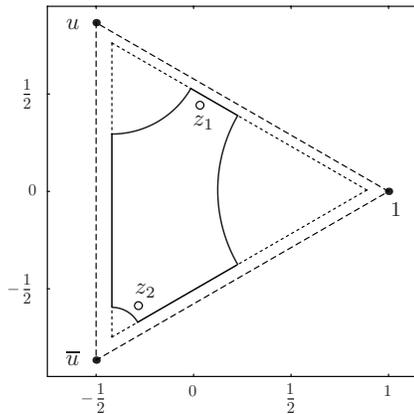


FIGURE 5.
 $2(z-1)^3(z-u)^2(z-\bar{u})$

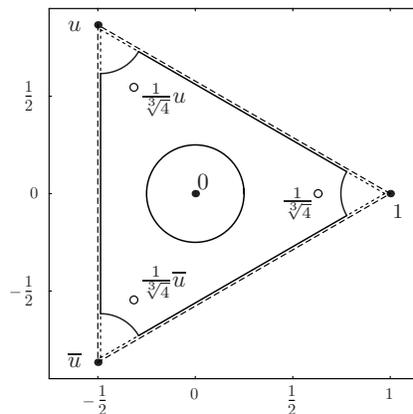


FIGURE 6. $z - z^4$

This translates in a shrunken Lucas polygon that is just the interval

$$[1/(n+2), 1 - 1/(n+2)].$$

However, the “Swiss-cheese” part of the theorem in this case gives something substantially better than the previous calculations, since shaving off the circles $D(0, 1/n)$ and $D(1, 1/n)$ from $[0, 1]$ leaves us with the interval $[1/n, 1 - 1/n]$, which is of course better (and optimal, since $1 - 1/n \in Z(p')$).

REFERENCES

- [1] P. Borwein, T. Erdélyi: *Polynomials and polynomial inequalities*, Graduate Texts in Mathematics, 161. Springer-Verlag, 1995. MR 97e:41001
- [2] B. Čurgus, V. Mascioni: *On the location of critical points of polynomials*, Proc. Amer. Math. Soc. 131 (2003), 253-264. MR 2003h:30006
- [3] D. Dimitrov: *A refinement of the Gauss-Lucas theorem*, Proc. Amer. Math. Soc. 126 (1998), 2065-2070. MR 98h:30005
- [4] P. Henrici: *Applied and computational complex analysis*, Vol. 1, John Wiley & Sons, 1988. MR 90d:30002
- [5] E. Hille: *Analytic function theory*, Vol. 1, Ginn and Company, Boston 1959. MR 21:6415
- [6] M. Marden: *The Location of the Zeros of the Derivative of a Polynomial*, American Mathematical Monthly, 42 (1935), 277-286.
- [7] M. Marden: *Geometry of polynomials*. Second edition reprinted with corrections, American Mathematical Society, Providence, RI, 1985. MR 37:1562
- [8] M. Mignotte: *Mathematics for computer algebra*. Springer-Verlag, 1992. MR 92i:68071
- [9] M. Mignotte, D. Ștefănescu: *Polynomials. An algorithmic approach*. Springer Series in Discrete Mathematics and Theoretical Computer Science. Springer-Verlag, Singapore, 1999. MR 2000e:12001
- [10] G. V. Milovanović, D. S. Mitrinović, Th. M. Rassias: *Topics in polynomials: extremal problems, inequalities, zeros*. World Scientific Publishing Co., 1994. MR 95m:30009
- [11] A. Turowicz: *Geometria zer wielomianów*. (Polish) [Geometry of zeros of polynomials] Państwowe Wydawnictwo Naukowe, Warsaw, 1967. MR 37:5372

- [12] J. L. Walsh: *The location of critical points of analytic and harmonic functions*, Amer. Math. Soc. Colloq. Publ., Vol. 34, 1950. MR 12:249d
- [13] R. Webster: *Convexity*, Oxford University Press, 1994. MR 98h:52001

DEPARTMENT OF MATHEMATICS, WESTERN WASHINGTON UNIVERSITY, BELLINGHAM, WASHINGTON 98225

E-mail address: `curgus@cc.wvu.edu`

DEPARTMENT OF MATHEMATICAL SCIENCES, BALL STATE UNIVERSITY, MUNCIE, INDIANA 47306-0490

E-mail address: `vdm@bsu-cs.bsu.edu`