A REFINEMENT OF THE GAUSS-LUCAS THEOREM

DIMITAR K. DIMITROV

ABSTRACT. The classical Gauss - Lucas Theorem states that all the critical points (zeros of the derivative) of a nonconstant polynomial p lie in the convex hull Ξ of the zeros of p. It is proved that, actually, a subdomain of Ξ contains the critical points of p.

1. Introduction and statement of results

Let

$$p(z)=\prod_{j=1}^{m}(z-z_j)^{k_j}, \quad \sum_{j=1}^{m}k_j=n, \tag{1}$$

be a polynomial of degree n whose zeros $z_1,\cdots,z_m$ are distinct and have multiplicities $k_1,\cdots,k_m$, respectively. Denote by Ξ the convex hull of $z_1,\cdots,z_m$. The Gauss - Lucas Theorem asserts that all the critical points of p lie in Ξ, and, furthermore, if the zeros of p are not collinear, no critical point of p lies on the boundary of Ξ unless it is a multiple zero of p. This classical result was implied in a note of Gauss dated 1836 and it was stated explicitly and proved by Lucas [1] in 1874. Many proofs of this theorem have been given but most of them duplicate Lucas' idea. It is based on a theorem of Gauss which provides a nice physical interpretation of the nontrivial critical points of a polynomial (the critical points which are not zeros of the polynomial) as the equilibrium points in a certain force field. The field is generated by particles placed at the zeros of the polynomial, the particles having masses equal to the multiplicity of the zeros and attracting with a force inversely proportional to the distance from the particle. We refer to Marden's book [2] for more information about Gauss-Lucas Theorem.

As it was pointed out by Marden [3, p.268], it is clear that the nontrivial critical points cannot be too close to any one zero since the force due to the particle at the zero would be relatively large. A partial quantitative result which provides an explicit form of the latter intuitive argument, is the following consequence of Exercise 4 on p. 92 of [2], which itself follows from Walsh's Two-circle Theorem [5]:

Theorem 1. *For any zero z_j of p, let $M_j:=\min_{l\neq j}|z_l-z_j|,\ j=1,\cdots,m$. Then p has no nontrivial critical point in*

$$\bigcup_{j=1}^{m}\left\{z:\quad |z-z_j|<\frac{k_j}{n}M_j\right\}.$$

1991 *Mathematics Subject Classification.* Primary 30C15, 26C10.

Key words and phrases. Nontrivial critical point of a polynomial.

Research supported by the Brazilian foudation CNPq under Grant 300645/95-3 and the Bulgarian Science Foundation under Grant MM-414.

We prove a further refinement of the Gauss-Lucas Theorem. Namely, we show that all the nontrivial critical points of a polynomial lie in a subdomain of the convex hull of the zeros of the polynomial, which does not contain certain neighbourhoods of the zeros. These neighbourhoods are larger than that specified in Theorem 1. Moreover, the main result allows certain neighbourhoods of the boundary of Ξ, which are free of critical points of p to be determined.

In what follows we suppose that p is defined by (1). With every zero z_j with multiplicity k_j we associate a closed circular region G_j in the complex plane which contains the points k_j/n and 1. Then, for each $l \neq j$, Ω_{jl} denotes the following affine transform of G_j:

$$\Omega_{jl} := z_j + (z_l - z_j)G_j.$$

Then we define

$$\Omega_j := \bigcup_{l \neq j} \Omega_{jl}.$$

The main result is:

Theorem 2. *Let, for every zero z_j of p, the region Ω_j be defined as above. Then every critical point of p which does not coincide with z_j lies in Ω_j. Moreover, if $\Omega_j, \quad j = 1, \cdots, m$, are all the regions associated with the distinct zeros of p, then every nontrivial critical point of p lies in*

$$\Omega(p) := \bigcap_{j=1}^{m} \Omega_j.$$

A consequence of the main result follows. In order to formulate it we need the notation

$$\Omega_{jl}^0 := \left\{ z : \quad |z - \left(\frac{n - k_j}{2n} z_j + \frac{n + k_j}{2n} z_l \right) | \leq \frac{n - k_j}{2n} |z_l - z_j| \right\}.$$

Corollary 1. *Every critical point of p which does not coincide with z_j lies in $\Omega_j^0 := \bigcup_{l \neq j} \Omega_{jl}^0$. Moreover, every nontrivial critical point of p lies in $\Omega^0(p) := \bigcap_{j=1}^{m} \Omega_j^0$.*

Note that the results are precise in the sense that there are polynomials for which all the nontrivial critical points lie on the boundaries of the corresponding regions. These polynomials are

$$p_{nk}(z) = z^k (z - 1)^{n-k}, \quad 1 \leq k \leq n - 1.$$

Indeed, since $z_1 = 0$ and $z_2 = 1$ then Ω_1 and Ω_2 are any closed circular domains which contain k/n and 0, and k/n and 1, respectively. Note that the only nontrivial critical point of p_{nk} is $\xi = k/n$. An application of Theorem 2 with regions Ω_1 and Ω_2 which touch at k/n yields the precise location of ξ. The discs $\Omega_1^0 = \Omega_{12}^0 = \left\{ z : \quad |z - \frac{n+k}{2n}| \leq \frac{n-k}{2n} \right\}$ and $\Omega_2^0 = \Omega_{21}^0 := \left\{ z : \quad |z - \frac{k}{2n}| \leq \frac{k}{2n} \right\}$ which appear in Corollary 1 are examples of such regions. Having in mind that $M_1 = M_2 = 1$, we see that the discs $\left\{ z : \quad |z| < \frac{k}{n} \right\}$ and $\left\{ z : \quad |z - 1| < \frac{n-k}{n} \right\}$ are the largest one to set in Theorem 1, namely, they are the largest discs centered at 0 and 1, respectively, which do not contain nontrivial critical points of p_{nk}.

2. Proofs

The basic tool in the proof of Theorem 2 is a result of Szegő [4] which is sometimes called Szegő's Composite Theorem. Let the polynomials A and B be defined by

$$A(z)=\sum_{j=0}^{n}\binom{n}{j}a_jz^j$$

and

$$B(z)=\sum_{j=0}^{n}\binom{n}{j}b_jz^j.$$

Then the polynomial

$$C(z)=\sum_{j=0}^{n}\binom{n}{j}a_jb_jz^j$$

is called composite of A and B.

Theorem 3. (Szegő's Composite Theorem) *Let all the zeros of A lie in the closed circular region D. Then every zero η of C can be represented in the form*

$$\eta=-w\ \beta_\mu,$$

where w is a point in D and β_μ is a zero of B.

Two simple technical lemmas follow.

Lemma 1. *For any positive integers $n\geq 2$ and k, $\ 1\leq k\leq n-1$, we have*

$$\sum_{\nu=0}^{n-k}\frac{k+\nu}{n}\binom{n-k}{\nu}z^\nu=(z+1)^{n-k-1}(z+k/n) \tag{2}$$

Proof. The coefficient of z^ν on the right-hand side of (2) is equal to

$$\begin{aligned}\frac{k}{n}\binom{n-k-1}{\nu}+\binom{n-k-1}{\nu-1} &= \frac{1}{n}\frac{(n-k-1)!}{\nu!(n-k-\nu)!}(n\nu+k(n-k-\nu))\\ &= \frac{k+\nu}{n}\binom{n-k}{\nu}.\end{aligned}$$

□

Let $n\geq 2$ and $k, 1\leq k\leq n-1$, and, let the polynomial P of degree n be of the form

$$P(z)=(z-a)^k\prod_{\nu=1}^{n-k}(z-\zeta_\nu).$$

Then

$$P'(z)=n(z-a)^{k-1}\prod_{\nu=1}^{n-k}(z-\xi_\nu).$$

The elementary symmetric functions of the variables $\alpha_1,\cdots,\alpha_{n-k}$ are denoted by $\sigma_0\equiv 1,\sigma_1(\alpha_1,\cdots,\alpha_{n-k}),\cdots,\sigma_{n-k}(\alpha_1,\cdots,\alpha_{n-k})$.

Lemma 2. *Let the polynomial P and its derivative P' be defined as above. Then, for every ν, $\ 0\leq\nu\leq n-k$, the identities*

$$\sigma_\nu(a-\xi_1,\cdots,a-\xi_{n-k})=\frac{n-\nu}{n}\ \sigma_\nu(a-\zeta_1,\cdots,a-\zeta_{n-k}) \tag{3}$$

hold.

Proof. Let $Q(z) := \prod_{\nu=1}^{n-k}(z-\zeta_\nu)$. Then $P(z) = (z-a)^k Q(z)$ and for every l, $0 \le l \le n-k$, we have

$$P^{(k+l)}(z) = \sum_{j=0}^{k+l} \left(\begin{array}{c} k+l \\ j \end{array} \right) \left[(z-a)^k\right]^{(j)} Q^{(k+l-j)}(z).$$

Since $\left[(z-a)^k\right]^{(j)}_{|z=a} = k!\delta_{kj}$, where δ_{kj} is the Kronecker delta, then

$$P^{(k+l)}(a) = \frac{(k+l)!}{l!} Q^{(l)}(a).$$

On the other hand, the identities $Q^{(l)}(z) = l!\ \sigma_{n-k-l}(z-\zeta_1, \cdots, z-\zeta_{n-k})$ hold. Hence

$$P^{(k+l)}(a) = (k+l)!\ \sigma_{n-k-l}(a-\zeta_1, \cdots, a-\zeta_{n-k}). \tag{4}$$

Similarly, denoting $R(z) := \prod_{\nu=1}^{n-k}(z-\xi_\nu)$, we get

$$P^{(k+l)}(z) = n \sum_{j=0}^{k+l-1} \left(\begin{array}{c} k+l-1 \\ j \end{array} \right) \left[(z-a)^k\right]^{(j)} R^{(k+l-1-j)}(z),$$

which yields

$$P^{(k+l)}(a) = n\frac{(k+l-1)!}{l!} R^{(l)}(a).$$

On using the identities $R^{(l)}(z) = l!\ \sigma_{n-k-l}(z-\xi_1, \cdots, z-\xi_{n-k})$ we obtain

$$P^{(k+l)}(a) = n(k+l-1)!\sigma_{n-k-l}(a-\xi_1, \cdots, a-\xi_{n-k}). \tag{5}$$

Now (4) and (5) yield

$$\sigma_{n-k-l}(a-\xi_1, \cdots, a-\xi_{n-k}) = \frac{k+l}{n}\ \sigma_{n-k-l}(a-\zeta_1, \cdots, a-\zeta_{n-k}),$$

which is equivalent to (3). $\square$

Proof of Theorem 2. Denote by $\zeta_1, \cdots, \zeta_{n-k}$ the zeros of p which are different from z_j. Then

$$p(z) = (z-z_j)^{k_j} \prod_{\nu=1}^{n-k_j} (z-\zeta_\nu).$$

The polynomial f of degree $n-k_j$ with leading coefficient one, whose zeros are $z_j - \zeta_\nu$, $\nu = 1, \cdots, n-k_j$, has the form

$$f(z) = \sum_{\nu=0}^{n-k_j} (-1)^{n-k_j-\nu} \sigma_{n-k_j-\nu}(z_j-\zeta_1, \cdots, z_j-\zeta_{n-k_j}) z^\nu.$$

Let

$$g(z) := \sum_{\nu=0}^{n-k_j} \frac{k_j+\nu}{n} \left(\begin{array}{c} n-k_j \\ \nu \end{array} \right) z^\nu.$$

It follows from Lemma 1 that $g(z) = (z+1)^{n-k_j-1}(z+k_j/n)$.

The polynomial h, which is the composite of g and f, is equal to

$$h(z) = \sum_{\nu=0}^{n-k_j} (-1)^{n-k_j-\nu} \sigma_{n-k_j-\nu}(z_j-\zeta_1, \cdots, z_j-\zeta_{n-k_j}) \frac{k_j+\nu}{n} z^\nu.$$

By Lemma 2 we have

$$\sigma_{n-k_j-\nu}(z_j-\xi_1,\cdots,z_j-\xi_{n-k_j})=\frac{k_j+\nu}{n}\sigma_{n-k_j-\nu}(z_j-\zeta_1,\cdots,z_j-\zeta_{n-k_j}),$$

where $\xi_1,\cdots,\xi_{n-k}$ are the critical points of p, which do not coincide with z_j. Hence

$$\begin{aligned} h(z) &= \sum_{\nu=0}^{n-k_j}(-1)^{n-k_j-\nu}\sigma_{n-k_j-\nu}(z_j-\xi_1,\cdots,z_j-\xi_{n-k_j})z^{\nu} \\ &= \prod_{\nu=1}^{n-k_j}(z-z_j+\xi_\nu). \end{aligned}$$

It follows from Theorem 3 that for every ν, $1\le\nu\le n-k_j$, $z_j-\xi_\nu$ can be represented in the form

$$z_j-\xi_\nu=-w\ (z_j-\zeta_\mu),$$

where μ, $1\le\mu\le n-k_j$, is an index and w is a point which belongs to a circular region which contains the zeros $-k_j/n$ and -1 of g. Hence $-w\in G_j$. Therefore for every critical point ξ_ν, which does not coincide with z_j, we have

$$\xi_\nu\in\bigcup_{l\neq j}(z_j+(z_l-z_j)G_j)\equiv\Omega_j.$$

This proves the first statement of the theorem. The second statement is an immediate consequence of the first one.

Proof of Corollary 1. For every j, $1\le j\le m$, we choose G_j to be the disc with diameter $[k_j/n,1]$,

$$G_j:=\left\{z:\quad |z-\frac{n+k_j}{2n}|\le\frac{n-k_j}{2n}\right\}.$$

Then for every $l\neq j$

$$\begin{aligned} z_j+(z_l-z_j)G_j &= \left\{z:\quad |z-\left(\frac{n-k_j}{2n}z_j+\frac{n+k_j}{2n}z_l\right)|\le\frac{n-k_j}{2n}|z_l-z_j|\right\} \\ &= \Omega_{jl}^0 \end{aligned}$$

and the statement of Corollary 1 follows immediately from Theorem 2.

Finally we show how Theorem 1 follows from Theorem 2.

Proof of Theorem 1. It has to be proved that p has no nontrivial critical point in neither of the discs $\left\{z:\quad |z-z_j|<\frac{k_j}{n}|z_l-z_j|\right\}$, $l\neq j$. One way to do this is to observe that $\Omega_j^0\cap\left\{z:\quad |z-z_j|<\frac{k_j}{n}M_j\right\}=\emptyset$. Another way is to choose the regions G_j in Theorem 2 to be the half-planes $\{z:\quad Re z\ge k_j/n\}$. It is easy to see that for this choice of G_j we have

$$\left(\bigcup_{l\neq j}(z_j+(z_l-z_j)G_j)\right)\bigcap\left\{z:\quad |z-z_j|<\frac{k_j}{n}M_j\right\}=\emptyset.$$

References

[1] F. Lucas, *Propriétés géométriques 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] M. Marden, *Geometry of Polynomials*, Amer.Math.Soc.Surveys, no.3, Providence, R.I., 1966.

[3] M. Marden, *Conjectures on the critical points of a polynomial*, Amer.Math.Monthly **90**(1983), 267-276.

[4] G. Szegő, *Bemerkungen zu einen Satz von J.H.Grace über die Wurzeln algebraischer Gleichungen*, Math.Z. **13**(1922), 28-55.

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

Departamento de Ciências de Computação e Estatística, IBILCE, Universidade Estadual Paulista, 15054-000 São José do Rio Preto, SP, Brazil

E-mail address: dimitrov@nimitz.dcce.ibilce.unesp.br