1

Roots of Polynomials

1.1 Inequalities for roots

1.1.1 The Fundamental Theorem of Algebra

In olden times, when algebraic theorems were scanty, the following statement received the title of the *Fundamental Theorem of Algebra*:

"*A given polynomial of degree n with complex coefficients has exactly n roots (multiplicities counted).*"

The first to formulate this statement was Alber de Girard in 1629, but he did not even try to prove it. The first to realize the necessity of proving the Fundamental Theorem of Algebra was d'Alembert. His proof (1746) was not, however, considered convincing. Euler (1749), Faunsenet (1759) and Lagrange (1771) offered their proofs but these proofs were not without blemishes, either.

The first to give a satisfactory proof of the Fundamental Theorem of Algebra was Gauss. He gave three different versions of the proof (1799, 1815 and 1816) and in 1845 he additionally published a refined version of his first proof.

For a review of the different proofs of the Fundamental Theorem of Algebra, see [Ti]. We confine ourselves to one proof. This proof is based on the following Rouché's theorem, which is of interest by itself.

Theorem 1.1.1 (Rouché). *Let f and g be polynomials, and γ a closed curve without self-intersections in the complex plane*[1]*. If*

$$\left|f(z) - g(z)\right| < \left|f(z)\right| + \left|g(z)\right| \tag{1}$$

for all $z \in \gamma$, then inside γ there is an equal number of roots of f and g (multiplicities counted).

[1] The plane $\mathbb{C}^1$ of complex variable.

V.V. Prasolov, *Polynomials*, Algorithms and Computation in Mathematics 11,
DOI 10.1007/978-3-642-03980-5_1,

Proof. In the complex plane, consider vector fields $v(z) = f(z)$ and $w(z) = g(z)$. From (1) it follows that at no point of γ are the vectors v and w directed opposite to each other. Recall that the *index* of the curve γ with respect to a vector field v is the number of revolutions of the vector $v(z)$ as it completely circumscribes the curve γ. (For a more detailed acquaintance with the properties of index we recommend Chapter 6 of [Pr2].) Consider the vector field

$$v_t = tv + (1-t)w.$$

Then $v_0 = w$ and $v_1 = v$. It is also clear that at every point $z \in \gamma$ the vector $v_t(z)$ is nonzero. This means that the index $\operatorname{ind}(t)$ of γ with respect to the vector field v_t is well defined. The integer $\operatorname{ind}(t)$ depends continuously on t, and hence $\operatorname{ind}(t) = \text{const}$. In particular, the indices of γ with respect to the vector fields v and w coincide.

Let the *index of the singular point* z_0 be defined as the index of the curve $|z - z_0| = \varepsilon$, where ε is sufficiently small. It is not difficult to show that the index of γ with respect to a vector field v is equal to the sum of indices of *singular* points, i.e., those at which $v(z) = 0$. For the vector field $v(z) = f(z)$, the index of the singular point z_0 is equal to the multiplicity of the root z_0 of f. Therefore the coincidence of the indices of γ with respect to vector fields $v(z) = f(z)$ and $w(z) = g(z)$ implies that, inside γ, the number of roots of f is equal to that of g. □

With the help of Rouché's theorem it is not only possible to prove the Fundamental Theorem of Algebra but also to estimate the absolute value of any root of the polynomial in question.

Theorem 1.1.2. *Let $f(z) = z^n + a_1 z^{n-1} + \dots + a_n$, where $a_i \in \mathbb{C}$. Then, inside the circle $|z| = 1 + \max\limits_i |a_i|$, there are exactly n roots of f (multiplicities counted).*

Proof. Let $a = \max\limits_i |a_i|$. Inside the circle considered, the polynomial $g(z) = z^n$ has root 0 of multiplicity n. Therefore it suffices to verify that, if $|z| = 1 + a$, then $\big|f(z) - g(z)\big| < \big|f(z)\big| + \big|g(z)\big|$. We will prove even that $\big|f(z) - g(z)\big| < \big|g(z)\big|$, i.e.,

$$|a_1 z^{n-1} + \dots + a_n| < |z|^n.$$

Clearly, if $|z| = 1 + a$, then

$$|a_1 z^{n-1} + \dots + a_n| \le a\big(|z|^{n-1} + \dots + 1\big) = a\frac{|z|^n - 1}{|z| - 1} = |z|^n - 1 < |z|^n. \quad \square$$

1.1.2 Cauchy's theorem

Here we discuss Cauchy's theorem on the roots of polynomials as well as its corollaries and generalizations.

Theorem 1.1.3 (Cauchy). *Let $f(x) = x^n - b_1x^{n-1} - \cdots - b_n$, where all the numbers b_i are non-negative and at least one of them is nonzero. The polynomial f has a unique (simple) positive root p and the absolute values of the other roots do not exceed p.*

Proof. Set
$$F(x) = -\frac{f(x)}{x^n} = \frac{b_1}{x} + \cdots + \frac{b_n}{x^n} - 1.$$
If $x \neq 0$, the equation $f(x) = 0$ is equivalent to the equation $F(x) = 0$. As x grows from 0 to $+\infty$ the function $F(x)$ strictly decreases from $+\infty$ to -1. Therefore, for $x > 0$, the function F vanishes at precisely one point, p. We have
$$-\frac{f'(p)}{p^n} = F'(p) = -\frac{b_1}{p^2} - \cdots - \frac{nb_n}{p^{n+1}} < 0.$$
Hence p is a simple root of f.

It remains to prove that if x_0 is a root of f, then $q = |x_0| \leq p$. Suppose that $q > p$. Then, since F is monotonic, it follows that $q > p$, i.e., $f(q) > 0$. On the other hand, the equality $x_0^n = b_1x_0^{n-1} + \cdots + b_n$ implies that
$$q^n \leq b_1q^{n-1} + \cdots + b_n,$$
i.e., $f(q) \leq 0$, which is a contradiction. $\square$

Remark. Cauchy's theorem is directly related to the Perron-Frobenius theorem on non-negative matrices (cf. [Wi1]).

The polynomial $x^{2n} - x^n - 1$ has n roots whose absolute values are equal to the value of the positive root of this polynomial. Therefore, in Cauchy's theorem, the estimate

the absolute values of the roots are $\leq p$

cannot, in general, be replaced by the estimate

the absolute values of the roots are $< p$.

Ostrovsky showed, nevertheless, that in a sufficiently general situation such a replacement is possible.

Theorem 1.1.4 (Ostrovsky). *Let $f(x) = x^n - b_1x^{n-1} - \cdots - b_n$, where all the numbers b_i are non-negative and at least one of them is nonzero.*

If the greatest common divisor of the indices of the positive coefficients b_i is equal to 1, then f has a unique positive root p and the absolute values of the other roots are $< p$.

Proof. Let only the coefficients $b_{k_1}, b_{k_2}, \dots, b_{k_m}$, where $k_1 < k_2 < \dots < k_m$, be positive. Since the greatest common divisor of $k_1, \dots, k_m$ is equal to 1, there exist integers $s_1, \dots, s_m$ such that $s_1 k_1 + \dots + s_m k_m = 1$. Consider again the function

$$F(x) = \frac{b_{k_1}}{x^{k_1}} + \dots + \frac{b_{k_m}}{x^{k_m}} - 1.$$

The equation $F(x) = 0$ has a unique positive solution p. Let x be any other (nonzero) root of f. Set $q = |x|$. Then

$$1 = \frac{b_{k_1}}{x^{k_1}} + \dots + \frac{b_{k_m}}{x^{k_m}} \leq \left|\frac{b_{k_1}}{x^{k_1}}\right| + \dots + \left|\frac{b_{k_m}}{x^{k_m}}\right| = \frac{b_{k_1}}{q^{k_1}} + \dots + \frac{b_{k_m}}{q^{k_m}},$$

i.e., $F(q) \geq 0$. We see that the equality $F(q) = 0$ is only possible if

$$\frac{b_{k_i}}{x^{k_i}} = \left|\frac{b_{k_i}}{x^{k_i}}\right| > 0 \text{ for all } i.$$

But in this case

$$\frac{b_{k_1}^{s_1} \cdot \ldots \cdot b_{k_m}^{s_m}}{x} = \left(\frac{b_{k_1}}{x^{k_1}}\right)^{s_1} \cdot \ldots \cdot \left(\frac{b_{k_m}}{x^{k_m}}\right)^{s_m} > 0,$$

i.e., $x > 0$. This contradicts the fact that $x \neq p$ and p is the only positive root of the equation $F(x) = 0$. Thus $F(q) > 0$. Therefore, since $F(x)$ is monotonic for positive x, it follows that $q < p$. □

The Cauchy-Ostrovsky theorem implies the following estimate of the absolute value of the roots of polynomials with positive coefficients.

Theorem 1.1.5. a) (Eneström-Kakeya) *If all the coefficients of the polynomial $g(x) = a_0 x^{n-1} + \dots + a_{n-1}$ are positive, then, for any root ξ of this polynomial, we have*

$$\min_{1 \leq i \leq n-1} \left\{\frac{a_i}{a_{i-1}}\right\} = \delta \leq |\xi| \leq \gamma = \max_{1 \leq i \leq n-1} \left\{\frac{a_i}{a_{i-1}}\right\}.$$

b) (Ostrovsky) *Let $\frac{a_k}{a_{k-1}} < \gamma$ for $k = k_1, \dots, k_m$. If the greatest common divisor of the numbers $n, k_1, \dots, k_m$ is equal to 1, then $|\xi| < \gamma$.*

Proof. Consider the polynomial

$$(x - \gamma) g(x) = a_0 x^n - (\gamma a_0 - a_1) x^{n-1} - \dots - (\gamma a_{n-2} - a_{n-1}) x - \gamma a_{n-1}.$$

By definition, $\gamma \geq \frac{a_i}{a_{i-1}}$, i.e., $\gamma a_{i-1} - a_i \geq 0$. Therefore, by Cauchy's theorem, γ is the only positive root of the polynomial $(x - \gamma) g(x)$ and the absolute values of the other roots of this polynomial are $\leq \gamma$.

If ξ is a root of g, then $\eta = \dfrac{1}{|\xi|}$ is a root of $a_{n-1}y^{n-1} + \cdots + a_0$. Hence

$$\frac{1}{|\xi|} = |\eta| = \max_{1 \le i \le n-1} \left\{ \frac{a_{i-1}}{a_i} \right\} = \frac{1}{\min\limits_{1 \le i \le n-1} \left\{ \frac{a_i}{a_{i-1}} \right\}},$$

i.e.,

$$|\xi| \ge \delta = \min_{1 \le i \le n-1} \left\{ \frac{a_i}{a_{i-1}} \right\}.$$

If condition b) is satisfied, the root γ of the polynomial $(x - \gamma)g(x)$ is strictly greater than the absolute values of the other roots of this polynomial. □

Remark. The Eneström-Kakeya theorem is also related to the Perron-Frobenius theorem, cf. [An2].

An essential generalization of the Eneström-Kakeya theorem is obtained in [Ga1]. However, the formulation of this generalization is rather cumbersome, and therefore we do not give it here.

1.1.3 Laguerre's theorem

Let $z_1, \dots, z_n \in \mathbb{C}$ be points of unit mass. The point $\zeta = \frac{1}{n}(z_1 + \cdots + z_n)$ is called the *center of mass* of $z_1, \dots, z_n$.

This notion can be generalized as follows. Perform a fractional-linear transformation w that sends z_0 to ∞, i.e.,

$$w(z) = \frac{a}{z - z_0} + b.$$

Let us find the center of mass of the images of $z_1, \dots, z_n$ and then apply the inverse transformation w^{-1}. Simple calculations show that the result does not depend on a and b, namely, we obtain the point

$$\zeta_{z_0} = z_0 + n \frac{1}{\frac{1}{z_1 - z_0} + \cdots + \frac{1}{z_n - z_0}} \tag{1}$$

which is called the *center of mass of* $z_1, \dots, z_n$ *with respect to* z_0.

Clearly,

the center of mass of $z_1, \dots, z_n$ *lies inside their convex hull.*

This statement easily generalizes to the case of the center of mass with respect to z_0. One only has to replace the lines that connect the points z_i and z_j by circles passing through z_i, z_j and z_0. The point z_0 corresponding to ∞ lies outside the convex hull.

Theorem 1.1.6. *Let $f(z) = (z - z_1) \cdot \ldots \cdot (z - z_n)$. Then the center of mass of the roots of f with respect to an arbitrary point z is given by the formula*

$$\zeta_z = z - n\frac{f(z)}{f'(z)}.$$

Proof. Clearly

$$\frac{f'(z)}{f(z)} = \frac{1}{z - z_1} + \cdots + \frac{1}{z - z_n}.$$

The desired statement follows directly from formula (1). □

Theorem 1.1.7 (Laguerre). *Let $f(z)$ be a polynomial of degree n and x its simple root. Then the center of mass of all the other roots of $f(z)$ with respect to x is the point*

$$X = x - 2(n-1)\frac{f'(x)}{f''(x)}.$$

Proof. Let $f(z) = (z - x)F(z)$. Then $f'(z) = F(z) + (z - x)F'(z)$ and $f''(z) = 2F'(z) + (z - x)F''(z)$. Therefore $f'(x) = F(x)$ and $f''(x) = 2F'(x)$. Applying the preceding theorem to the polynomial F of degree $n - 1$, and point $z = x$, we obtain the desired statement. □

Theorem 1.1.8 (Laguerre). *Let $f(z)$ be a polynomial of degree n and*

$$X(z) = z - 2(n-1)\frac{f'(z)}{f''(z)}.$$

Let the circle (or line) C pass through a simple root z_1 of f and the other roots of f belong to one of the two domains into which C divides the plane. Then $X(z_1)$ also belongs to the same domain.

Proof. In the case of the "usual" center of mass, the circle C corresponds to the line such that all the roots of $f(z)$, except z_1, lie on one side of it. The center of mass of these roots lies on the same side of this line. □

Corollary. *Let z_1 be one of the simple roots of f with the maximal absolute value. Then $|X(z_1)| \leq |z_1|$, i.e.,*

$$\left| z_1 - 2(n-1)\frac{f'(z_1)}{f''(z_1)} \right| \leq |z_1|.$$

Proof. All the roots of f lie in the disk $\left\{ z \in \mathbb{C} \,\middle|\, |z| \leq |z_1| \right\}$, and therefore $X(z_1)$ also belongs to this disk. □

Theorem 1.1.9. *Let f be a polynomial with real coefficients and define*

$$\zeta_z = z - n\frac{f(z)}{f'(z)}.$$

All the roots of f are real if and only if $\operatorname{Im} z \cdot \operatorname{Im} \zeta_z < 0$ *for any* $z \in \mathbb{C} \setminus \mathbb{R}$.

Proof. Suppose first that all the roots of f are real. Let $\operatorname{Im} z = a > 0$. The line consisting of the points with the imaginary part ε, where $0 < \varepsilon < a$, separates the point z from all the roots of f since they belong to the real axis. Therefore $\operatorname{Im} \zeta_z \leq \varepsilon$. In the limit as $\varepsilon \to 0$, we obtain $\operatorname{Im} \zeta_z \leq 0$.

It is easy to verify that it is impossible to have $\operatorname{Im} \zeta_z = 0$. Indeed, let $\zeta_z \in \mathbb{R}$. Consider a circle passing through z and tangent to the real axis at ζ_z. Slightly jiggling this circle we can construct a circle on one side of which lie the points z and ζ_z, and on the other side lie all the roots of f. If $\operatorname{Im} z = a < 0$, the arguments are similar.

Now suppose that $\operatorname{Im} z \cdot \operatorname{Im} \zeta_z < 0$ for all $z \in \mathbb{C} \setminus \mathbb{R}$. Let z_1 be a root of f such that $\operatorname{Im}(z_1) \neq 0$. Then $\lim\limits_{z \to z_1} \zeta_z = z_1$, and therefore $\operatorname{Im} z_1 \cdot \operatorname{Im} \zeta_{z_1} > 0$. □

Our presentation of Laguerre's theory is based on the paper [Gr], see also [Pol].

1.1.4 Apolar polynomials

Let $f(z)$ be a polynomial of degree n and ζ a fixed number or ∞. The function

$$A_\zeta f(z) = \begin{cases} (\zeta - z)f'(z) + nf(z) & \text{if } \zeta \neq \infty; \\ f'(z) & \text{if } \zeta = \infty \end{cases}$$

is called *the derivative of* $f(z)$ with respect to point ζ. It is easy to verify that, if

$$f(z) = \sum_{k=0}^{n} \binom{n}{k} a_k z^k, \tag{1}$$

then

$$\frac{1}{n} f'(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} a_{k+1} z^k. \tag{$*$}$$

Therefore

$$\frac{1}{n} A_\zeta f(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} (a_k + a_{k+1}\zeta) z^k. \tag{2}$$

Let $z_1, \dots, z_n$ be the roots of the polynomial (1), and let $\zeta_1, \dots, \zeta_n$ be the roots of the polynomial

$$g(z) = \sum_{k=0}^{n} \binom{n}{k} b_k z^k. \tag{3}$$

Formula (2) implies that

$$\frac{1}{n!} A_{\zeta_1} A_{\zeta_2} \cdots A_{\zeta_n} f(z) = a_0 + a_1\sigma_1 + a_2\sigma_2 + \dots + a_n\sigma_n,$$

where

$$\sigma_1 = \zeta_1 + \zeta_2 + \cdots + \zeta_n = -\binom{n}{1}\frac{b_{n-1}}{b_n},$$
$$\sigma_2 = \zeta_1\zeta_2 + \cdots + \zeta_{n-1}\zeta_n = \binom{n}{2}\frac{b_{n-2}}{b_n},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\sigma_n = \zeta_1 \cdot \ldots \cdot \zeta_n = (-1)^n\frac{b_0}{b_n}.$$

Hence the equality $A_{\zeta_1}A_{\zeta_2}\cdots A_{\zeta_n}f(z) = 0$ is equivalent to the equality

$$a_0b_n - \binom{n}{1}a_1b_{n-1} + \binom{n}{2}a_2b_{n-2} + \cdots + (-1)^n a_nb_0 = 0. \tag{4}$$

The polynomials f and g given by (1) and (3) and whose coefficients are related via (4) are said to be *apolar.*

A *circular domain* is either the inner or the exterior part of a disk or the half plane.

Theorem 1.1.10 (J. H. Grace, 1902). *Let f and g be apolar polynomials. If all the roots of f belong to a circular domain K, then at least one of the roots of g also belongs to K.*

Proof. We will need the following auxiliary statement.

Lemma 1.1.11. *Let all the roots $z_1, \ldots, z_n$ of $f(z)$ lie inside the circular domain K and let ζ lie outside K. Then all the roots of $A_\zeta f(z)$ lie inside K.*

Proof. Observe first that, if w_i is a root of the polynomial $A_\zeta f(z)$, then ζ is the center of mass of the roots of $f(z)$ with respect to w_i. Indeed, if $\zeta \neq \infty$, then we can express the equality $A_\zeta f(w_i) = 0$ in the form

$$(\zeta - w_i)f'(w_i) + nf(w_i) = 0, \quad \text{i.e.,} \quad \zeta = w_i - n\frac{f(w_i)}{f'(w_i)}.$$

If $\zeta = \infty$, then $f'(w_i) = A_\zeta f(w_i) = 0$, and hence

$$\sum_{j=1}^{n}\frac{1}{z_j - w_i} = \frac{f'(w_i)}{f(w_i)} = 0.$$

Therefore the center of mass of the points $z_1, \ldots, z_n$ with respect to w_i is situated at

$$w_i + \frac{1}{\sum_j \frac{1}{z_j - w_i}} = \infty.$$

Now it is clear that point w_i cannot lie outside K. Indeed, if w_i were situated outside K, then the center of mass of $z_1, \ldots, z_n$ with respect to w_i would be inside K. However, this contradicts the fact that ζ lies outside K. $\square$

With the help of Lemma 1.1.11, Theorem 1.1.10 is proved as follows. Suppose that all the roots $\zeta_1, \dots, \zeta_n$ of g lie outside K. Consider the polynomial $A_{\zeta_2} \cdots A_{\zeta_n} f(z)$. Its degree is equal to 1, i.e., it is of the form $c(z-k)$. Lemma 1.1.11 implies that $k \in K$. Since f and g are apolar polynomials, it follows that $A_{\zeta_1}(z-k) = 0$. On the other hand, the direct calculation of the derivative shows that $A_{\zeta_1}(z-k) = \zeta_1 - k$. Therefore $k = \zeta_1 \notin K$ and we have a contradiction. □

Every polynomial f has a whole family of polynomials apolar to it. Having selected a convenient apolar polynomial we can, thanks to Grace's theorem, prove that f possesses a root in a given circular domain. Sometimes for the same goal it is convenient to use Lemma 1.1.11 directly.

Example 1. The polynomial

$$f(z) = 1 - z + cz^n, \quad \text{where} \quad c \in \mathbb{C},$$

possesses a root in the disk $|z-1| \le 1$.

Proof. The polynomials

$$f(z) = 1 + \binom{n}{1}\frac{-1}{n}z + cz^n \text{ and } g(z) = z^n + \binom{n}{1}b_{n-1}z^{n-1} + \cdots + b_0$$

are apolar if

$$1 - n\left(\frac{-1}{n}\right)b_{n-1} + cb_0 = 0, \quad \text{i.e.,} \quad 1 + b_{n-1} + cb_0 = 0.$$

Now let $\zeta_k = 1 - \exp(2\pi i k/n)$ for $k = 1, \dots, n$, and take $g(z)$ to be

$$g(z) = \prod (z - \zeta_k) = z^n + \binom{n}{1}b_{n-1}z^{n-1} + \cdots + b_0.$$

Then

$$b_{n-1} = -1 \text{ and } b_0 = \pm \prod \zeta_k = 0.$$

Therefore the polynomials $f(z)$ and $g(z)$ are apolar. Since all the roots of g lie in the disk $|z-1| \le 1$, at least one of the roots of f lies in this disk. □

Example 2. The polynomial $1 - z + c_1 z^{n_1} + \cdots + c_k z^{n_k}$, where $1 < n_1 < n_2 < \cdots < n_k$, has at least one root in the disk

$$|z| \le \frac{1}{\left(1 - \frac{1}{n_1}\right) \cdot \ldots \cdot \left(1 - \frac{1}{n_k}\right)}.$$

Proof. Let us start with the polynomial $f(z) = 1 - z + c_1 z^{n_1}$. Suppose on the contrary that all its roots lie in the domain $|z| > \frac{n_1}{n_1-1}$. Then by Lemma 1.1.11 the roots of the polynomial

$$A_0 f(z) = n_1 - (n_1 - 1)z$$

also lie in the domain $|z| > \frac{n_1}{n_1-1}$. But the root of $A_0 f(z)$ is equal to $\frac{n_1}{n_1-1}$ and we have a contradiction.

For the polynomial $f(z) = 1 - z + c_1 z^{n_1} + \cdots + c_k z^{n_k}$, we use induction on k. Consider the polynomial

$$A_0 f(z) = n_k - (n_k - 1)z + c_1(n_k - n_1)z^{n_1} + \cdots + c_{k-1}(n_k - n_{k-1})z^{n_{k-1}}.$$

In this polynomial, replace z by $\frac{n_k}{n_k-1}w$. By the induction hypothesis, the roots of the polynomial obtained lie in the disk

$$|w| \leq \frac{n_1}{n_1 - 1} \cdot \frac{n_2}{n_2 - 1} \cdot \ldots \cdot \frac{n_{k-1}}{n_{k-1} - 1},$$

and hence the roots of $A_0 f(z)$ lie in the disk

$$|z| \leq \frac{n_1}{n_1 - 1} \cdot \frac{n_2}{n_2 - 1} \cdot \ldots \cdot \frac{n_k}{n_k - 1}.$$

Therefore the hypothesis that all the roots of $f(z)$ lie outside the disk leads to a contradiction. □

Let $f(z) = \sum\limits_{i=1}^{n} \binom{n}{i} a_i z^i$ and $g(z) = \sum\limits_{i=1}^{n} \binom{n}{i} b_i z^i$. The polynomial

$$h(z) = \sum_{i=1}^{n} \binom{n}{i} a_i b_i z^i$$

is called the *composition* of f and g.

Theorem 1.1.12 (Szegö). *Let f and g be polynomials of degree n, and let all the roots of f lie in a circular domain K. Then every root of the composition h of f and g is of the form $-\zeta_i k$, where ζ_i is a root of g and $k \in K$.*

Proof. Let γ be a root of h, i.e., $\sum\limits_{i=1}^{n} \binom{n}{i} a_i b_i \gamma^i = 0$. Then the polynomials $f(z)$ and $G(z) = z^n g(-\gamma z^{-1})$ are apolar. Therefore, by Grace's theorem, one of the roots of $G(z)$ lies in K. Let, for example, $g(-\gamma k^{-1}) = 0$, where $k \in K$. Then $-\gamma k^{-1} = \zeta_i$, where ζ_i is a root of g. □

For polynomials whose degrees are not necessarily equal, there is the following analogue of Grace's theorem.

Theorem 1.1.13 ([Az]). *Let* $f(z) = \sum\limits_{i=1}^{n} \binom{n}{i} a_i z^i$ *and* $g(z) = \sum\limits_{i=1}^{m} \binom{m}{i} b_i z^i$ *be polynomials with* $m \le n$*. Let the coefficients of* f *and* g *be related as follows:*

$$\binom{m}{0} a_0 b_m - \binom{m}{1} a_1 b_{m-1} + \cdots + (-1)^m \binom{m}{m} a_m b_0 = 0. \tag{5}$$

Then the following statements hold:

a) *If all the roots of* $g(z)$ *belong to the disk* $|z| \le r$*, then at least one of the roots of* $f(z)$ *also belongs to this disk;*

b) *If all the roots of* $f(z)$ *lie outside the disk* $|z| \le r$*, then at least one of the roots of* $g(z)$ *also lies outside this disk.*

Proof. [Ru] a) Relation (5) is invariant with respect to the change of z to rz in f and g, and therefore we may assume that $r = 1$. Suppose on the contrary that all the roots of $f(z)$ lie in the domain $|z| > 1$. Then all the roots of the polynomial $z^n f(\frac{1}{z})$ lie in the domain $|z| < 1$. Therefore, from the Gauss-Lucas theorem (Theorem 1.2.1 on p. 13), it follows that all the roots of the polynomial

$$f_1(z) = D^{(n-m)} \left(z^n f\left(\frac{1}{z}\right) \right) = n(n-1) \cdot \ldots \cdot (m+1) \sum_{i=0}^{m} \binom{m}{i} a_i z^{m-i}$$

lie in the domain $|z| < 1$. Therefore all the roots of the polynomial

$$f_2(z) = z^m \sum_{i=0}^{m} \binom{m}{i} a_i \left(\frac{1}{z}\right)^{m-i} = \sum_{i=0}^{m} \binom{m}{i} a_i z^i$$

lie in the domain $|z| > 1$.

Relation (5) means that the polynomials f_2 and g are apolar. Since all the roots of f_2 lie in the circular domain $|z| > 1$, it follows from Grace's theorem that at least one of the roots of g also lies in this domain, and we have a contradiction.

b) All the roots of f_2 lie in the domain $|z| \ge 1$, hence, it follows from Grace's theorem that at least one of the roots of g also lies in this domain. □

1.1.5 The Routh-Hurwitz problem

In various problems on stability one has to investigate whether all the roots of a given polynomial belong to the left half-plane (i.e., whether the real parts of the roots are negative). The polynomials with this property are said to be *stable*. The Routh-Hurwitz problem is

how to find out directly by looking at the coefficients of the polynomial whether it is stable or not.

Several different solutions of the problem are known (see, e.g., [Po2]). We will confine ourselves with one simple criterion given in [St3].

First, we observe that it suffices to consider the case of polynomials with real coefficients. Indeed, if $p(z) = \sum a_n z^n$ is a polynomial with complex coefficients we can consider the polynomial

$$p^*(z) = p(z)\overline{p(\overline{z})} = \left(\sum a_n z^n\right)\left(\sum \overline{a_n} z^n\right).$$

Clearly, the real parts of the roots of $\overline{p(\overline{z})}$ are the same as those of $p(z)$. Moreover, the coefficients of $p^*(z)$ are symmetric with respect to a_n and $\overline{a_n}$. This means that the coefficients of p^* are invariant under conjugation, that is, they are real.

Theorem 1.1.14. *Let $p(z) = z^n + a_1 z^{n-1} + \cdots + a_n$ be a polynomial with real coefficients; let $q(z) = z^m + b_1 z^{m-1} + \cdots + b_m$, where $m = \frac{1}{2}n(n-1)$, be the polynomial whose roots are all the sums of pairs of the roots of p. The polynomial p is stable if and only if all the coefficients of the polynomials p and q are positive.*

Proof. Suppose that p is stable. To a negative root α of p there corresponds the factor $z-\alpha$ with positive coefficients. To a pair of conjugate roots with the negative real part there corresponds the factor

$$(z-\alpha-i\beta)(z-\alpha+i\beta) = z^2 - 2\alpha z + \alpha^2 + \beta^2$$

with positive coefficients. Thus all the coefficients of p are positive.

The complex roots of q fall into the pairs of conjugate roots because the coefficients of q are real. Further, the real parts of all the roots of q are negative. The same arguments as for p show that all the coefficients of q are positive.

Next, let all the coefficients of p and q be positive. In this case, all the real roots of p and q are negative. Therefore, if α is a real root of p, then $\alpha < 0$, and, if $\alpha \pm i\beta$ is a pair of complex conjugate roots of p, then $2\alpha = (\alpha+i\beta)+(\alpha-i\beta)$ is a root of q; hence $2\alpha < 0$. $\square$

1.2 The roots of a given polynomial and of its derivative

1.2.1 The Gauss-Lucas theorem

In 1836, Gauss showed that all the roots of P', distinct from the multiple roots of the polynomial P itself, serve as the points of equilibrium for the field of forces created by identical particles placed at the roots of P (provided that r particles are located at the root of multiplicity r) if each particle creates an attractive force inversely proportional to the distance to this particle. From this theorem of Gauss it is easy to deduce Theorem 1.2.1 given below. Gauss himself did not mention this. The first to formulate and prove Theorem 1.2.1 was a French engineer F. Lucas in 1874. Therefore Theorem 1.2.1 is often referred to as the *Gauss-Lucas theorem.*

Theorem 1.2.1 (Gauss-Lucas). *The roots of P' belong to the convex hull of the roots of the polynomial P itself.*

Proof. Let $P(z) = (z - z_1) \cdot \ldots \cdot (z - z_n)$. It is easy to verify that

$$\frac{P'(z)}{P(z)} = \frac{1}{z - z_1} + \cdots + \frac{1}{z - z_n}. \tag{1}$$

Suppose that $P'(w) = 0$, $P(w) \neq 0$ and suppose on the contrary that w does not belong to the convex hull of the points $z_1, \ldots, z_n$. Then one can draw a line through w that does not intersect the convex hull of $z_1, \ldots, z_n$. Therefore the vectors $w - z_1, \ldots, w - z_n$ lie in one half-plane determined by this line. Hence the vectors

$$\frac{1}{w - z_1}, \ldots, \frac{1}{w - z_n}$$

also lie in one half-plane, since $\dfrac{1}{z} = \dfrac{\overline{z}}{|z|^2}$. Hence,

$$\frac{P'(w)}{P(w)} = \frac{1}{w - z_1} + \cdots + \frac{1}{w - z_n} \neq 0.$$

This is a contradiction, and hence w belongs to the convex hull of the roots of P. $\square$

Relation (1) allows one to prove the following properties of the roots of P' for any polynomial P with real roots.

Theorem 1.2.2 ([An1]). *Let*

$$P(z) = (z - x_1) \cdot \ldots \cdot (z - x_n), \text{ where } x_1 < \cdots < x_n.$$

If some root x_i is replaced by $x_i' \in (x_i, x_{i+1})$, then all the roots of P' increase their value.

Proof. Let $z_1 < z_2 < \cdots < z_{n-1}$ be the roots of P', and let $x_1, \ldots, x_n$ be the roots of P. Let $z_1' < z_2' < \cdots < z_{n-1}'$ be the roots of Q' and let $x_1' = x_1, \ldots, x_{i-1}' = x_{i-1}, x_i', x_{i+1}' = x_{i+1}, \ldots, x_n' = x_n$ be the roots of Q. For the roots z_k and z_k', the relation (1) takes the form

$$\sum_{i=1}^{n} \frac{1}{z_k - x_i} = 0, \quad \sum_{i=1}^{n} \frac{1}{z_k' - x_i'} = 0. \tag{2}$$

Suppose that the statement of the theorem is false, i.e., $z_k' < z_k$ for some k. Then $z_k' - x_i' < z_k - x_i$. Observe that the differences $z_k' - x_i'$ and $z_k - x_i$ are of the same sign. Indeed,

$$z_j < x_i, \; z_j' < x_i' \text{ for } j \leq i - 1 \text{ and } z_j > x_i, \; z_j' > x_i' \text{ for } j \geq i.$$

Hence, $\frac{1}{z_k - x_i} < \frac{1}{z_k' - x_i'}$ for all $i = 1, \ldots, n$. But in this case relations (2) cannot hold simultaneously. $\square$

1.2.2 The roots of the derivative and the focal points of an ellipse

The roots of the derivative of a cubic polynomial have the following interesting geometric interpretation.

Theorem 1.2.3 (van der Berg, [Be2]). *Let the roots of a cubic polynomial P form the vertices of a triangle ABC in the complex plane. Then the roots of P' are at the focal points of the ellipse tangent to the sides of $\triangle ABC$ at their midpoints.*

First proof. Observe first of all that if $Q(z) = P(z - z_0)$, then $Q'(z) = P'(z - z_0)$. Therefore we can take any point for the origin.

We can represent any affine transformation of the plane as a composition of an isometry, a homothety, and a transformation of the form $(x, y) \mapsto (x, y \cos\alpha)$ in a Cartesian coordinate system. Therefore we may assume that the triangle ABC is obtained from the equilateral triangle with vertices w, εw and $\varepsilon^2 w$, where $|w| = 1$ and $\varepsilon = \exp\left(\frac{2\pi i}{3}\right)$, under the transformation

$$z \mapsto \frac{z + \overline{z}}{2} + \frac{z - \overline{z}}{2}\cos\alpha = z\cos^2\frac{\alpha}{2} + \overline{z}\sin^2\frac{\alpha}{2}. \tag{1}$$

Then the semi-axes a and b of the ellipse considered are equal to $\frac{1}{2}$ and $\frac{1}{2}\cos\alpha$; the distance between its focal points F_1 and F_2 is equal to $\sqrt{a^2 - b^2} = \frac{1}{2}\sin\alpha$. Under the dilation with coefficient

$$\left(\frac{1}{2}\sin\alpha\right)^{-1} = \left(\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\right)^{-1}.$$

points F_1 and F_2 transform into $(\pm 1, 0)$. The composition of transformation (1) and this dilation amounts to the transformation

$$z \mapsto z\cot\frac{\alpha}{2} + \overline{z}\tan\frac{\alpha}{2}$$

Set $a = w\cot\frac{\alpha}{2}$. Then the polynomial with roots A, B, and C is of the form

$$P(x) = \left(x - a - \frac{1}{a}\right)\left(x - a\varepsilon - \frac{1}{a\varepsilon}\right)\left(x - a\varepsilon^2 - \frac{1}{a\varepsilon^2}\right)$$

It is easy to verify that $P'(x) = 3x^2 + 3\varepsilon + 3\overline{\varepsilon} = 3x^2 - 3$, and therefore the roots of P' are ± 1. $\square$

Second proof. [Sc5] Let $\varepsilon = \exp\left(\frac{2\pi i}{3}\right)$ and let z_0, z_1, z_2 be the roots of the polynomial P considered. Select numbers $\zeta_0, \zeta_1, \zeta_2$ so that

$$z_0 = \zeta_0 + \zeta_1 + \zeta_2, \quad z_1 = \zeta_0 + \zeta_1\varepsilon + \zeta_2\varepsilon^2, \quad z_2 = \zeta_0 + \zeta_1\varepsilon^2 + \zeta_2\varepsilon, \tag{2}$$

i.e.,

$$3\zeta_0 = z_0 + z_1 + z_2, \quad 3\zeta_1 = z_0 + z_1\varepsilon^2 + z_2\varepsilon, \quad 3\zeta_2 = z_0 + z_1\varepsilon + z_2\varepsilon^2.$$

In what follows we assume that $z_0 + z_1 + z_2 = 0$, i.e., $\zeta_0 = 0$.

It is easy to verify that the curve $\zeta_1 e^{i\varphi} + \zeta_2 e^{-i\varphi}$, where $0 \le \varphi \le 2\pi$, is an ellipse whose semi-axes are directed along the bisectors of the exterior and interior angles of the angle $\angle\zeta_1 O \zeta_2$, where O is the origin, and the lengths of the semi-axes are equal to $|\zeta_1| + |\zeta_2|$ and $||\zeta_1| - |\zeta_2||$. Indeed, the curve considered is the image of the unit circle under the map $z \mapsto \zeta_1 z + \zeta_2 \overline{z}$. Further, if $\zeta_1 = |\zeta_1| e^{i\alpha}$ and $\zeta_2 = |\zeta_2| e^{i\beta}$, then

$$\zeta_1 e^{i\varphi} + \zeta_2 e^{-i\varphi} = |\zeta_1| e^{i(\varphi+\alpha)} + |\zeta_2| e^{i(\beta-\varphi)}.$$

The absolute value of this expression attains its maximum at $\varphi = \frac{\alpha+\beta}{2} + k\pi$ and its minimum at $\varphi = \frac{\alpha+\beta}{2} + \frac{\pi}{2} + k\pi$. These values of φ correspond precisely to the directions of the bisectors indicated.

The focal points f_1 and f_2 of the ellipse $\zeta_1 e^{i\varphi} + \zeta_2 e^{-i\varphi}$ lie on the line corresponding to the angle $\varphi = \frac{\alpha+\beta}{2}$, i.e., $\frac{f_1 f_2}{\zeta_1 \zeta_2}$ is a positive number. Further, the square of the distance of the focal point to the center of the ellipse is equal to the difference of the squares of the semi-axes, i.e., it is equal to

$$\left(|\zeta_1| + |\zeta_2|\right)^2 - \left(|\zeta_1| - |\zeta_2|\right)^2 = 4|\zeta_1\zeta_2|.$$

Hence $f_1 f_2 = 4\zeta_1\zeta_2$.

Relations (2) for $\zeta_0 = 0$ show that the vertices z_0, z_1, z_2 of the triangle considered lie on the ellipse $\zeta_1 e^{i\varphi} + \zeta_2 e^{-i\varphi}$ and the mid-points of its sides lie on the ellipse $\frac{1}{2}\left(\zeta_1 e^{i\varphi} + \zeta_2 e^{-i\varphi}\right)$. The mid-point of a chord of the first ellipse lies on the second ellipse only if this chord is tangent to the second ellipse. Therefore we have to prove that the focal points of the ellipse $\frac{1}{2}\left(\zeta_1 e^{i\varphi} + \zeta_2 e^{-i\varphi}\right)$ coincide with the roots of the derivative of the polynomial $P = (z - z_0)(z - z_1)(z - z_2)$. The focal points of the ellipse satisfy the equation $z^2 - \zeta_1\zeta_2 = 0$, and the roots of P' satisfy

$$3z^2 + z_0 z_1 + z_0 z_2 + z_1 z_2 = 0, \quad \text{i.e.,} \quad 3(z^2 - \zeta_1\zeta_2) = 0. \quad \square$$

1.2.3 Localization of the roots of the derivative

Jensen's disks

Let f be a polynomial with real coefficients. For every pair of conjugate roots z and $\overline{z}$ of f, the disk with diameter[1] $z\overline{z}$ is called a *Jensen's disk* of f.

Theorem 1.2.4 (Jensen). *Any non-real root of f' lies inside or on the boundary of one of the Jensen's disks of f.*

[1] We mean that z and $\overline{z}$ are the endpoints of a diameter of this disk.

Proof. Let $z_1, \dots, z_n$ be the roots of f. Then

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

Let us show first of all that if z lies outside Jensen's disk with diameter $z_p z_q$, then

$$\operatorname{sgn} \operatorname{Im} \left(\frac{1}{z - z_p} + \frac{1}{z - z_q} \right) = - \operatorname{sgn} \operatorname{Im} z. \tag{2}$$

Indeed,

$$\frac{1}{z - a - bi} + \frac{1}{z - a + bi} = \frac{2(z-a)\left((\overline{z} - a)^2 + b^2\right)}{|(z-a)^2 + b^2|^2}$$

and

$$\operatorname{Im} \left((\overline{z} - a)|z - a|^2 + (z - a)b^2 \right) = \left(b^2 - |z - a|^2 \right) \operatorname{Im} z.$$

Let us show now that if $z \notin \mathbb{R}$ and $z_j = a \in \mathbb{R}$, then

$$\operatorname{sgn} \operatorname{Im} \left(\frac{1}{z - z_j} \right) = - \operatorname{sgn} \operatorname{Im} z. \tag{3}$$

Indeed,

$$\frac{1}{z - a} - \frac{1}{\overline{z} - a} = \frac{\overline{z} - z}{|z - a|^2} = \frac{-2 \operatorname{Im} z}{|z - a|^2}.$$

Formulas (1), (2), (3) imply that if point $z \notin \mathbb{R}$ lies outside all the Jensen's disks, then

$$\operatorname{sgn} \operatorname{Im} \frac{f'(z)}{f(z)} = - \operatorname{sgn} \operatorname{Im} z \neq 0.$$

Hence $f'(z) \neq 0$, i.e., z is not a root of f'. □

As a refinement of Jensen's theorem, we prove the following estimate for the number of the roots of the derivative whose real parts belongs to a given segment.

Theorem 1.2.5 (Walsh). *Let $I = [\alpha, \beta]$, and let K be the union of I and Jensen's disks intersecting I. If K contains k roots of a polynomial $f(z)$, then the number of the roots of $f'(z)$ that lie in K is between $k - 1$ and $k + 1$.*

Proof. Let C be the boundary of the smallest rectangle whose sides are parallel to the coordinate axes and which contain K. Consider the restriction to C of the map $z \mapsto e^{i\varphi}$, where $\varphi = \arg \frac{f'(z)}{f(z)}$. Formulas (1), (2) and (3) imply that the image of the part of C that lies in the upper half-plane lies on the half-circle $|z| = 1$, $\operatorname{Im} z \leq 0$, whereas the image of the part of C that lies in the lower half-plane lies on the half-circle $|z| = 1$, $\operatorname{Im} z \geq 0$. Therefore the number of revolutions of the image of C around the origin is equal to either 0 or ± 1. This means that the indices of C with respect to the vector fields $f(z)$ and $f'(z)$ either coincide or differ by ± 1, i.e., the total numbers of the zeros of functions f and f' lying inside C either coincide or differ by ± 1. □

Walsh's theorem

Theorem 1.2.6 (Walsh). *Let the roots of the polynomials f_1 and f_2 lie in the disks K_1 and K_2 with radii r_1 and r_2 and centers at points c_1 and c_2, respectively. Then every root of the derivative of $f = f_1 f_2$ lie either in K_1, or in K_2, or in the disk of radius $\dfrac{n_2 r_1 + n_1 r_2}{n_1 + n_2}$ centered at $\dfrac{n_2 c_1 + n_1 c_2}{n_1 + n_2}$, where $n_1 = \deg f_1$ and $n_2 = \deg f_2$.*

Proof. Let z be the root of f lying outside K_1 and K_2. Then

$$f_1'(z) f_2(z) + f_1(z) f_2'(z) = 0;$$

moreover, $f_1(z), f_2(z), f_1'(z), f_2'(z)$ are nonzero.

Consider ζ_1 and ζ_2, the centers of mass of the roots of f_1 and f_2 with respect to z, respectively. By Theorem 1.1.6

$$\zeta_1 = z - n_1 \frac{f_1(z)}{f_1'(z)}, \quad \zeta_2 = z - n_2 \frac{f_2(z)}{f_2'(z)}.$$

Hence

$$n_2 \zeta_1 + n_1 \zeta_2 = (n_1 + n_2) z - n_1 n_2 \left(\frac{f_1(z)}{f_1'(z)} + \frac{f_2(z)}{f_2'(z)} \right) = (n_1 + n_2) z,$$

i.e., $z = \dfrac{n_2 \zeta_1 + n_1 \zeta_2}{n_1 + n_2}$. Since all the roots of f_i lie in K_i, it follows that $\zeta_i \in K_i$. It remains to observe that if points ζ_1 and ζ_2 of mass n_2 and n_1 lie in disks K_1 and K_2, respectively, then their center of mass z lies in the disk K. □

The Grace-Heawood theorem

Theorem 1.2.7 (J. H. Grace, 1902; P. J. Heawood, 1907). *If z_1 and z_2 are distinct roots of a polynomial f of degree n, then the disk $|z - c| \le r$, where $c = \frac{1}{2}(z_1 + z_2)$ and $r = \dfrac{|z_1 - z_2|}{2} \cot\left(\dfrac{\pi}{n}\right)$, contains at least one root of f'.*

Proof. Let[1] $f'(z) = \sum\limits_{k=0}^{n-1} \binom{n-1}{k} a_k z^k$. Then

$$0 = f(z_2) - f(z_1) = \int_{z_1}^{z_2} f'(z)\, dz = \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} a_k b_{n-1-k},$$

where the coefficients $b_0, \ldots, b_{n-1}$ depend only on z_1 and z_2 and not on the coefficients $a_0, \ldots, a_{n-1}$. Therefore, given z_1 and z_2, we can construct a polynomial $g(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} b_k z^k$ apolar to $f'(z)$.

[1] This expression of f' differs by a factor of $\frac{1}{n}$ from formula (∗) in sec. 1.1.4.

To obtain an explicit formula for g, set $a_k = (-1)^k x^{n-1-k}$, i.e., consider $h(z) = (x-z)^{n-1}$. In this case

$$g(x) = \sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-1-k} b_{n-1-k} =$$
$$\int_{z_1}^{z_2} (x-1)^{n-1} dz = \frac{(x-z_1)^n - (x-z_2)^n}{n}.$$

The roots of g are of the form

$$\zeta_k = \frac{z_1+z_2}{2} + i\frac{z_1-z_2}{2} \cot \frac{k\pi}{n} \quad \text{for } k = 1, 2, \ldots, n-1$$

and all of them lie on the boundary of the disk considered. Therefore, by Theorem 1.1.10 (see p. 8), the disk $|z-c| \leq r$ contains at least one root of f'. □

In [Ma7], there are several other theorems on localization of the roots of the derivative.

1.2.4 The Sendov-Ilieff conjecture

In 1962, the Bulgarian mathematician B. Sendov made the following conjecture often ascribed to another Bulgarians mathematician, L. Ilieff:

"*Let* $P(z)$ *be a polynomial* ($\deg P \geq 2$) *all of whose roots lie in the disk* $|z| \leq 1$. *If* z_0 *is one of the roots of* $P(z)$, *then the disk* $|z - z_0| \leq 1$ *contains at least one root of* $P'(z)$".

This conjecture is proved for all polynomials of degree ≤ 5 and several particular polynomials (see, e.g., [Sc4]).

We confine ourselves to the proof of the conjecture for polynomials of the form

$$P(z) = (z-z_0)^{n_0}(z-z_1)^{n_1}(z-z_2)^{n_2}.$$

This proof is given in [Co2].

The case when $n = n_0 + n_1 + n_2 \geq 4$ is much the simplest. In this case we have to prove that if $|z_i| \leq 1$ for $i = 0, 1, 2$, then the polynomial

$$P'(z) = n(z-z_0)^{n_0-1}(z-z_1)^{n_1-1}(z-z_2)^{n_2-1}(z-w_1)(z-w_2) \qquad (1)$$

has a root lying in the disk $|z - z_0| \leq 1$. If $n_0 > 1$, then z_0 is such a root. We assume therefore that $n_0 = 1$. Let us express $P(z)$ in the form $P(z) = (z-z_0)Q(z)$. It is clear that

$$P'(z_0) = Q(z_0) = (z_0-z_1)^{n_1}(z_0-z_2)^{n_2}. \qquad (2)$$

It follows from (1) and (2) that

$$n(z_0 - w_1)(z_0 - w_2) = (z_0 - z_1)(z_0 - z_2). \tag{3}$$

Taking into account that $|z_0 - z_1| \leq |z_0| + |z_1| = 2$ and $|z_0 - z_2| \leq 2$, we obtain

$$|z_0 - w_1| \cdot |z_0 - w_2| \leq \frac{4}{n} \leq 1,$$

and hence either $|z_0 - w_1| \leq 1$ or $|z_0 - w_2| \leq 1$.

It remains to consider the case when $n_0 = n_1 = n_2 = 1$. For this we need the following auxiliary statement which we will formulate more generally than is needed for this proof.

Lemma. *Let $P(z)$ be a polynomial of degree n, where $n \geq 2$. If*

$$|P''(z_0)| \geq (n-1)\,|P'(z_0)|,$$

then at least one of the roots of P' lies inside the disk $|z - z_0| \leq 1$.

Proof. Let $w_1, w_2, \ldots, w_{n-1}$ be the roots of P'. We may assume that the highest coefficient of P is equal to 1. In this case $P'(z) = n \prod_{j=1}^{n-1} (z - w_j)$. If $P'(z) \neq 0$ we may take the logarithm of both sides and differentiate. This gives

$$\frac{P''(z)}{P'(z)} = \sum_{j=1}^{n-1} \frac{1}{z - w_j}.$$

By the hypothesis z_0 is a simple root of P, i.e., $P'(z_0) \neq 0$. Suppose that $|z_0 - w_j| > 1$ for $j = 1, \ldots, n-1$. Then the inequality $|P''(z_0)| \geq (n-1)\,|P'(z_0)|$ implies that

$$n - 1 \leq \left| \frac{P''(z_0)}{P'(z_0)} \right| \leq \sum_{j=1}^{n-1} \frac{1}{|z_0 - w_j|} < n - 1,$$

and we have a contradiction. □

Now let us consider directly the polynomial

$$P(z) = (z - z_0)(z - z_1)(z - z_2) = (z - z_0)Q(z).$$

Clearly

$$\frac{P''(z)}{P'(z)} = 2\frac{Q'(z)}{Q(z)} = 2\left(\frac{1}{z_0 - z_1} + \frac{1}{z_0 - z_2} \right) = \frac{2(2z_0 - z_1 - z_2)}{(z_0 - z_1)(z_0 - z_2)}.$$

Now consider the triangle ABC with vertices $A = z_0$, $B = z_1$, $C = z_2$. Obviously $|z_0 - z_1| = c$, $|z_0 - z_2| = b$ and $|2z_0 - z_1 - z_2| = 2m_a$, where m_a is the length of the median drawn from A. By Lemma the Sendov-Ilieff conjecture holds if $4m_a \geq 2bc$.

By the hypothesis, $b \le 2$ and $c \le 2$, and hence $2m_a \ge bc$ holds both for $m_a \ge b$ and for $m_a \ge c$. It remains to consider the case when $m_a < b$ and $m_a < c$.

Relation (3) shows that the Sendov-Ilieff conjecture holds if $bc \le 3$. Therefore we may assume that $bc > 3$. In this case

$$b^2 + c^2 = (b - c)^2 + 2bc > 6,$$

and therefore $b^2 + c^2 - a^2 > 6 - 4 > 0$, i.e., $\angle A < 90°$. The inequalities $b > m_a$ and $c > m_a$ imply that $\angle C < 90°$ and $\angle B < 90°$, and hence the triangle ABC is acute.

Let R be the radius of its circumscribed circle, h_a the length of the altitude from A. Then $\dfrac{c}{h_a} = \sin B = \dfrac{2R}{b}$, i.e., $bc = 2Rh_a \le 2Rm_a$. To obtain the inequality desired, $bc \le 2m_a$, it remains therefore to prove that $R \le 1$. The acute triangle ABC lies inside the unit circle $|z| = 1$. If the circumscribed circle S of the triangle ABC lies inside the unit circle, the inequality $R \le 1$ is obvious. Let now S and the unit circle have a common chord. Since ABC is acute, this chord subtends an acute angle φ whose vertex coincides with one of the vertices of the triangle ABC. The same chord subtends the angles ψ and $180° - \psi$, where $\psi \le 90°$, whose vertices lie on the unit circle. Moreover, $\psi \le \varphi$. The inequalities $\psi \le \varphi < 90° < 180° - \psi$ imply that $R \le 1$.

1.2.5 Polynomials whose roots coincide with the roots of their derivatives

In the paper [Ya] it was stated that if P and Q are *monic* polynomials (i.e., their highest coefficients are equal to 1) and the sets of roots of P and Q coincide, and the sets of roots of the polynomials P' and Q' also coincide, then $P^m = Q^n$ for certain positive integers m and n. Later certain gaps were discovered in the proof of this statement and soon a counterexample was constructed in [Ro2]. The construction of this counterexample is rather complicated. We advise the interested reader to turn directly to [Ro2].

Concerning properties of polynomials whose roots coincide with the roots of the derivatives see also [Do1].

1.3 The resultant and the discriminant

1.3.1 The resultant

Consider polynomials $f(x) = \sum\limits_{i=0}^{n} a_i x^{n-i}$ and $g(x) = \sum\limits_{i=0}^{m} b_i x^{m-i}$, where $a_0 \ne 0$ and $b_0 \ne 0$. Over an algebraically closed field, f and g have a common divisor if and only if they have a common root. If the field is not algebraically closed, then the common divisor could be a polynomial without roots.

The existence of a common divisor of f and g is equivalent, as one can show, to the existence of polynomials p and q such that $fq = gp$, where $\deg p \leq n-1$ and $\deg q \leq m-1$. Indeed, let $f = hp$ and $g = hq$. Then $fq = hpq = gp$. Suppose now that $fq = gp$, where $\deg q \leq \deg g - 1$. If f and g do not have a common divisor, then q divides g: a contradiction.

Let $q = u_0x^{m-1} + \dots + u_{m-1}$ and $p = v_0x^{n-1} + \dots + v_{n-1}$. The equality $fq = gp$ can be expressed as a system of equations:

$$\begin{aligned} a_0u_0 &= b_0v_0, \\ a_1u_0 + a_0u_1 &= b_1v_0 + b_0v_1, \\ a_2u_0 + a_1u_1 + a_0u_2 &= b_2v_0 + b_1v_1 + b_0v_2, \\ \dots\dots\dots\dots\dots\dots &\dots\dots\dots\dots \end{aligned}$$

The polynomials f and g have a common root if and only if this system has a nonzero solution $(u_0, u_1, \dots, v_0, v_1, \dots)$. If, for example, $m = 3$ and $n = 2$, the determinant of this system of equations is of the form

$$\begin{vmatrix} a_0 & 0 & 0 & -b_0 & 0 \\ a_1 & a_0 & 0 & -b_1 & -b_0 \\ a_2 & a_1 & a_0 & -b_2 & -b_1 \\ 0 & a_2 & a_1 & -b_3 & -b_2 \\ 0 & 0 & a_2 & 0 & -b_3 \end{vmatrix} = \pm \begin{vmatrix} a_0 & a_1 & a_2 & 0 & 0 \\ 0 & a_0 & a_1 & a_2 & 0 \\ 0 & 0 & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & b_3 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 \end{vmatrix} = \pm \det S(f, g).$$

The matrix

$$S(f, g) = \begin{pmatrix} a_0 & a_1 & a_2 & 0 & 0 \\ 0 & a_0 & a_1 & a_2 & 0 \\ 0 & 0 & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & b_3 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 \end{pmatrix}$$

is called the *Sylvester matrix* of the polynomials f and g. The determinant of $S(f, g)$ is called the *resultant* of f and g and is denoted by $R(f, g)$. Clearly, $R(f, g)$ is a homogeneous polynomial of degree m with respect to indeterminates a_i and of degree n with respect to indeterminates b_j. The polynomials f and g have a common divisor if and only if the determinant of the system considered vanishes, i.e., $R(f, g) = 0$.

The resultant has many different applications. For example, given polynomial relations $P(x, z) = 0$ and $Q(y, z) = 0$ we can, with the help of the resultant, obtain a polynomial relation of the form $R(x, y) = 0$, i.e., eliminate z. Indeed, consider the given polynomials $P(x, z)$ and $Q(y, z)$ as polynomials in z regarding x and y as constants. Then the vanishing of the resultant of these polynomials is exactly the relation desired $R(x, y) = 0$.

The resultant also allows one to reduce the solution of the system of algebraic equations to the search for roots of polynomials. Indeed, let $P(x_0, y_0) = 0$ and $Q(x_0, y_0) = 0$. Consider $P(x, y)$ and $Q(x, y)$ as polynomials in y. For $x = x_0$, they have a common root y_0. Therefore their resultant $R(x)$ vanishes at $x = x_0$.

Theorem 1.3.1. *Let x_i be the roots of f, and y_j the roots of g. Then*

$$R(f,g) = a_0^m b_0^n \prod (x_i - y_j) = a_0^m \prod g(x_i) = b_0^n \prod f(y_j).$$

Proof. Since $f(x) = a_0(x - x_1) \cdot \ldots \cdot (x - x_n)$, it follows that

$$a_k = \pm a_0 \sigma_k(x_1, \ldots, x_n),$$

where σ_k is an elementary symmetric function. Similarly,

$$b_k = \pm b_0 \sigma_k(y_1, \ldots, y_m).$$

The resultant is a homogeneous polynomial of degree m with respect to indeterminates a_i and of degree n with respect to the b_j. Hence

$$R(f,g) = a_0^m b_0^n P(x_1, \ldots, x_n, y_1, \ldots, y_m),$$

where P is a symmetric polynomial in $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ vanishing at $x_i = y_j$. The formula

$$x_i^k = (x_i - y_j)x_i^{k-1} + y_j x_i^{k-1}$$

shows that

$$P(x_1, \ldots, y_m) = (x_i - y_j)Q(x_1, \ldots, y_m) + U(x_1, \ldots, \widehat{x}_i, \ldots, y_m).$$

Substituting $x_i = y_j$ into this formula we see that U is the zero polynomial. Similar arguments show that P is divisible by $S = a_0^m b_0^n \prod (x_i - y_j)$.

Since $g(x) = b_0 \prod\limits_{j=1}^{m} (x - y_j)$, we have $\prod\limits_{i=1}^{n} g(x_i) = b_0^n \prod\limits_{i,j} (x_i - y_j)$, and therefore

$$S = a_0^m \prod_{i=1}^{n} g(x_i) = a_0^m \prod_{i=1}^{n} (b_0 x_i^m + b_1 x_i^{m-1} + \cdots + b_m)$$

is a homogeneous polynomial of degree n in indeterminates $b_0, \ldots, b_m$. For indeterminates $a_0, \ldots, a_n$, the arguments are similar. It is also clear that the symmetric polynomial $a_0^m \prod_{i=1}^n (b_0 x_i^m + b_1 x_i^{m-1} + \cdots + b_m)$ is a polynomial in $a_0, \ldots, a_n, b_0, \ldots, b_m$. Hence $R(f,g) = R(a_0, \ldots, b_m) = \lambda S$, where λ is a number which does not depend on the a_i and b_i.

On the other hand, the coefficient of $\prod x_i^m$ in $a_0^m b_0^n P(x_1, \ldots, y_m)$ and S is equal to $a_0^m b_0^n$, hence, $\lambda = 1$. □

Corollary 1. $R(g,f) = (-1)^{\deg f \deg g} R(f,g)$.

Corollary 2. *If $f = gq + r$, then*

$$R(f,g) = b_0^{\deg f - \deg r} R(r,g),$$

where b_0 is the leading coefficient of g.

Proof. Let y_j be the roots of g. Then $f(y_j) = r(y_j)$. It remains use that $R(f,g) = b_0^{\deg f} \prod f(y_j)$ and $R(r,g) = b_0^{\deg r} \prod f(y_j)$. □

Corollary 3. $R(f, gh) = R(f, g)R(f, h)$

Proof. Let x_i be the roots of f and a_0 its leading coefficient. Then

$$\begin{aligned} R(f, gh) &= a_0^{\deg g + \deg h} \prod g(x_i)h(x_i), \\ R(f, g) &= a_0^{\deg g} \prod g(x_i), \\ R(f, h) &= a_0^{\deg h} \prod h(x_i). \ \square \end{aligned}$$

Theorem 1.3.2. *Let* $f(x) = \sum_{i=0}^{n} a_i x^{n-i}$ *and* $g(x) = \sum_{i=0}^{m} b_i x^{m-i}$. *Then there exist polynomials* φ *and* ψ *with integer coefficients in indeterminates* $a_0, \dots, a_n$, $b_0, \dots, b_m$ *and* x *for which the identity*

$$\varphi(x, a, b)f(x) + \psi(x, a, b)g(x) = R(f, g)$$

holds.

Proof. Let $c_0, \dots, c_{n+m-1}$ be the columns of the Sylvester matrix $S(f, g)$ and $y_k = x^{m+n-k-1}$. Then

$$y_0 c_0 + \dots + y_{n+m-1} c_{n+m-1} = c,$$

where c is the column vector

$$\left(x^{m-1} f(x), \dots, f(x), x^{n-1} g(x), \dots, g(x)\right)^T.$$

Consider $y_0, \dots, y_{n+m-1}$ as a system of linear equations for $y_0, \dots, y_{n+m-1}$ and make use of Cramer's rule in order to find y_{n+m-1}. We obtain

$$y_{n+m-1} \det(c_0, \dots, c_{n+m-1}) = \det(c_0, \dots, c_{n+m-2}, c). \tag{1}$$

It remains to notice that $y_{n+m-1} = 1$, $\det(c_0, \dots, c_{n+m-1}) = R(f, g)$ and the determinant on the right-hand side of (1) can be represented in the form desired, i.e., as $\varphi(x, a, b)f(x) + \psi(x, a, b)g(x)$. □

1.3.2 The discriminant

Let $x_1, \dots, x_n$ be the roots of the polynomial $f(x) = a_0 x^n + \dots + a_n$, where $a_0 \neq 0$. The quantity

$$D(f) = a_0^{2n-2} \prod_{i<j} (a_i - a_j)^2$$

is called the *discriminant* of f.

Theorem 1.3.3. *$R(f, f') = \pm a_0 D(f)$.*

Proof. By Theorem 1.3.1 we have $R(f, f') = a_0^{n-1} \prod_i f'(x_i)$. It is easy to verify that $f'(x_i) = a_0 \prod_{j \neq i} (x_j - x_i)$. Therefore

$$R(f, f') = a_0^{2n-1} \prod_{j \neq i} (x_j - x_i) = \pm a_0^{2n-1} \prod_{i<j} (x_j - x_i)^2. \ \square$$

Remark. It is not difficult to show that

$$R(f, f') = -R(f', f) = (-1)^{n(n-1)/2} a_0 D(f).$$

Corollary. *The discriminant of f is a polynomial in $a_0, \dots, a_n$ with integer coefficients.*

Theorem 1.3.4. *Let f, g, and h be monic polynomials. Then*

$$D(fg) = D(f)D(g)R^2(f, g)$$
$$D(fgh) = D(f)D(g)D(h)R^2(f, g)R^2(g, h)R^2(h, f).$$

Proof. Let $x_1, \dots, x_n$ be the roots of f, and $y_1, \dots, y_m$ the roots of g. Then

$$D(fg) = \prod (x_i - x_j)^2 \prod (y_i - y_j)^2 \prod (x_i - y_j)^2 = D(f)D(g)R^2(f, g).$$

The second formula is proved similarly. $\square$

Theorem 1.3.5. *Let f be a real polynomial of degree n without real roots. Then* $\operatorname{sgn} D(f) = (-1)^{n/2}$.

Proof. Making use of the factorization

$$f(x) = a_0(x - x_1) \cdot \ldots \cdot (x - x_n)$$

it is easy to verify that

$$D\left((x - a)f(x)\right) = D\left(f(x)\right)\left(f(a)\right)^2.$$

Let a and $\overline{a}$ be a pair of conjugate roots of f, i.e., $f(x) = (x-a)(x-\overline{a})g(x)$. Then

$$D\left(f(x)\right) = D\left(g(x)\right)(a - \overline{a})^2 \left(f(a)f(\overline{a})\right)^2.$$

Clearly, $\operatorname{sgn}(a - \overline{a})^2 = -1$ and $\left(f(a)f(\overline{a})\right)^2 = |f(a)|^4 > 0$. Therefore $\operatorname{sgn} D(f) = -\operatorname{sgn} D(g)$. Now it is easy to obtain the statement required by induction on n. $\square$

Theorem 1.3.6. *Let $f(x) = x^n + a_1 x^{n-1} + \dots + a_n$ be a polynomial with integer coefficients. Then its discriminant $D(f)$ is equal to either $4k$ or $4k+1$, where k is an integer.*

Proof. Let $x_1, \dots, x_n$ be the roots of f. Then

$$D(f) = \delta^2(f), \quad \text{where } \delta(f) = \prod_{i<j}(x_i - x_j).$$

Consider an auxiliary polynomial $\delta_1(f) = \prod_{i<j}(x_i + x_j)$. Clearly, $\delta_1(f)$ is a symmetric function of the roots of f, and hence $\delta_1(f)$ is an integer. Moreover,

$$\delta_1^2(f) - \delta^2(f) = \prod_{i<j}\left((x_i - x_j)^2 + 4x_i x_j\right) - \prod_{i<j}(x_i - x_j)^2 = 4U(x_1, \dots, x_n),$$

where U is a symmetric polynomial in $x_1, \dots, x_n$ with integer coefficients. Therefore $D(f) = \delta_1^2(f) + 4k_1$, where k_1 is an integer. It is also clear that $\delta_1^2(f) = 4k_2$ or $4k_2 + 1$. □

1.3.3 Computing certain resultants and discriminants

In this section we give several examples on how to compute resultants and discriminants.

Example 1.3.7. $D(x^n + a) = (-1)^{n(n-1)/2} n^n a^{n-1}$.

Proof. Let us make use of the fact that

$$D(f) = (-1)^{n(n-1)/2} R(f, f') = (-1)^{n(n-1)/2} \prod_{i=1}^{n} f'(x_i),$$

where $x_1, \dots, x_n$ are the roots of f. In our case $f'(x) = nx^{n-1}$ and $\prod x_i = (-1)^n a$, and therefore $\prod x_i^{n-1} = (-1)^{n(n-1)} a^{n-1} = a^{n-1}$. □

Example 1.3.8. Let $\varphi(x) = x^{n-1} + \cdots + 1$. Then $D(\varphi) = (-1)^{(n-1)(n-2)/2} n^{n-2}$.

Proof. Since $(x-1)\varphi(x) = x^n - 1$, it follows that

$$D(\varphi)\left(\varphi(1)\right)^2 = D\left((x-1)\varphi(x)\right) = D(x^n - 1) = (-1)^{(n-1)(n-2)/2} n^n.$$

It remains to observe that $\varphi(1) = n$. □

Example 1.3.9. Let $f_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$. Then

$$D(n! f_n) = (-1)^{n(n-1)/2} (n!)^n.$$

Proof. The polynomial $g_n = n! f_n$ is monic, and hence

$$D(g) = (-1)^{n(n-1)/2} R(g, g') = (-1)^{n(n-1)/2} \prod_{i=1}^{n} g'(\alpha_i),$$

where $\alpha_1, \dots, \alpha_n$ are the roots of f_n. Clearly,

$$g'(\alpha_i) = n! f_n'(\alpha_i) = n! f_{n-1}(\alpha_i) = n! \left(f_n(\alpha_i) - \frac{\alpha_i^n}{n!} \right) = -\alpha_i^n.$$

Therefore

$$D(g) = (-1)^{n(n-1)/2} \prod_{i=1}^{n} (-\alpha_i^n).$$

It remains to observe that $\prod \alpha_i = (-1)^n g(0) = (-1)^n n!$. □

Example 1.3.10. Let $d = (r, s)$, $r_1 = \dfrac{r}{d}$ and $s_1 = \dfrac{s}{d}$. Then

$$R(x^r - a, x^s - b) = (-1)^s \left(a^{s_1} - b^{r_1}\right)^d.$$

Proof. The relation $R(g, f) = (-1)^{\deg f \deg g} R(f, g)$ shows that if the desired statement holds for a pair (r, s), then it also holds for a pair (s, r). Indeed, $(-1)^{rs+d+r} = (-1)^s$. We may therefore assume that $r \geq s$.

For $s = 0$, the statement is obvious. If $s > 0$, then having divided $x^r - a$ by $x^s - b$ we get the residue $bx^{r-s} - a$. Hence

$$\begin{aligned} R(x^r - a, x^s - b) &= R(bx^{r-s} - a, x^s - b) = \\ &= R(b, x^s - b) R\left(x^{r-s} - \frac{a}{b}, x^s - b\right) = \\ &= b^s R\left(x^{r-s} - \frac{a}{b}, x^s - b\right). \end{aligned}$$

It is easy to see that if $R(x^{r-s} - \frac{a}{b}, x^s - b) = (-1)^s \left((\frac{a}{b})^{s_1} - b^{r_1 - s_1}\right)$, then

$$R(x^r - a, x^s - b) = (-1)^s \left(a^{s_1} - b^{r_1}\right)^d.$$

It remains to use induction on $r + s$. □

Example 1.3.11. Let $n > k > 0$, $d = (n, k)$, $n_1 = \dfrac{n}{d}$ and $k_1 = \dfrac{k}{d}$. Then

$$\begin{aligned} &D(x^n + ax^k + b) = \\ &(-1)^{n(n-1)/2} b^{k-1} \left(n^{n_1} b^{n_1 - k_1} + (-1)^{n_1+1} (n-k)^{n_1 - k_1} k^{k_1} a^{n_1}\right)^d. \end{aligned}$$

Proof. [Sw] The formula $D(f) = (-1)^{n(n-1)/2} R(f, f')$ gives

$$\begin{aligned} D(x^n + ax^k + b) &= (-1)^{n(n-1)/2} R(x^n + ax^k + b, nx^{n-1} + kax^{k-1}) = \\ &= (-1)^{n(n-1)/2} n^n R(x^n + ax^k + b, x^{n-1} + n^{-1} kax^{k-1}). \end{aligned}$$

Using the fact that

$$R(f, x^m g) = R(f, x^m) R(f, g) = (f(0))^m R(f, g),$$

we obtain

$$D(x^n + ax^k + b) = (-1)^{n(n-1)/2} n^n b^{k-1} R(x^n + ax^k + b, x^{n-k} + n^{-1}ka).$$

The residue after the division of $x^n + ax^k + b$ by $x^{n-k} + n^{-1}ka$ is equal to $a(1 - n^{-1}k)x^k + b$, and hence

$$R(x^n + ax^k + b, x^{n-k} + n^{-1}ka) = R\left(a(1 - n^{-1}k)x^k + b, x^{n-k} + n^{-1}ka\right).$$

The resultant of a pair of two binomials is computed in Example 1.3.10. □

1.4 Separation of roots

Here we discuss various theorems which allow us to compute, or at least estimate from above, the number of real roots of a polynomial on a given segment (a, b). Formulations of such theorems often use the notion the *number of sign changes* in the sequence $a_0, a_1, \dots, a_n$, where $a_0 a_n \neq 0$. This number is determined as follows: all the zero terms of the sequence considered are deleted and, for the remaining non-zero terms, one counts the number of pairs of neighboring terms of different sign.

1.4.1 The Fourier–Budan theorem

Theorem 1.4.1 (Fourier–Budan). *Let $N(x)$ be the number of sign changes in the sequence $f(x)$, $f'(x)$, ..., $f^{(n)}(x)$, where f is a polynomial of degree n. Then the number of roots of f (multiplicities counted) between a and b, where $f(a) \neq 0$, $f(b) \neq 0$ and $a < b$, does not exceed $N(a) - N(b)$. Moreover, the number of roots can differ from $N(a) - N(b)$ by an even number only.*

Proof. Let x be a point which moves along the segment $[a, b]$ from a to b. The number $N(x)$ varies only if x passes through a root of the polynomial $f^{(m)}$ for some $m \leq n$.

Consider first the case when x passes through a root x_0 of multiplicity r of $f(x)$. In a neighborhood of x_0, the polynomials $f(x), f'(x), \dots, f^{(r)}(x)$ behave approximately as

$$(x - x_0)^r g(x_0), \quad (x - x_0)^{r-1} r g(x_0), \quad \dots, \quad r! g(x_0),$$

respectively. Therefore, for $x < x_0$, there are r sign changes in this sequence and for $x > x_0$ there are no sign changes (assuming that x is sufficiently close to x_0).

Now suppose that x passes through a root x_0 of multiplicity r of $f^{(m)}$; let x_0 be not a root of $f^{(m-1)}$. (Of course, x_0 can be a root of f as well, as it can be not a root of f.) We have to prove that under the passage through x_0 the number of sign changes in the sequence $f^{(m-1)}(x)$, $f^{(m)}(x)$, ..., $f^{(m+r)}(x)$

changes by a non-negative even integer. Indeed, in a vicinity of x_0 these polynomials behave approximately as

$$F(x_0),\ (x-x_0)^r G(x_0),\ (x-x_0)^{r-1} rG(x_0),\ \ldots,\ r!G(x_0). \tag{1}$$

Excluding $F(x_0)$, we see that the remaining system has exactly r sign changes for $x < x_0$ and no sign changes for $x > x_0$. Concerning the first two terms, $F(x_0)$ and $(x-x_0)^r G(x_0)$, of the sequence (1) we see that if r is even, then the number of sign changes is the same for $x < x_0$ and $x > x_0$ whereas if r is odd, then the number of sign changes for $x < x_0$ is by 1 greater or less than for $x > x_0$ (depending whether $F(x_0)$ and $G(x_0)$ have the same sign or the opposite sign). Thus, for r even, the difference in the number of sign changes is equal to r and, for r odd, the difference of the number of sign changes is equal to $r \pm 1$. In both these cases this difference is even and non-negative. □

Corollary 1. (The Descartes Rule) *The number of positive roots of the polynomial* $f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n$ *does not exceed the number of sign changes in the sequence* $a_0, a_1, \ldots, a_n$.

Proof. Since $f^{(r)}(0) = r!a_{n-r}$, it follows that $N(0)$ coincides with the number of sign changes in the sequence of coefficients of f. It is also clear that $N(+\infty) = 0$. □

Remark. Jacobi showed that the Descartes Rule can be used also to estimate the number of roots between α and β. To this end one should make the change of variables $y = \dfrac{x-\alpha}{\beta - x}$, i.e., set $x = \dfrac{\alpha + \beta y}{1+y}$, and consider the polynomial

$$(1+y)^n f\left(\frac{\alpha+\beta y}{1+y}\right) = b_0 y^n + b_1 y^{n-1} + \cdots + b_n.$$

The Descartes Rule applied to this polynomial yields an estimate of the number of roots between α and β. Indeed, y varies from 0 to ∞, as x varies from α to β.

Corollary 2. (de Gua) *If the polynomial lacks* $2m$ *consecutive terms (i.e., the coefficients of these terms vanish), then this polynomial has no less than* $2m$ *imaginary roots. If* $2m+1$ *consecutive terms are missing, then if they are between terms of different signs, the polynomial has no less than* $2m$ *imaginary roots, whereas if the missing terms are between terms of the same sign the polynomial has no less than* $2m+2$ *imaginary roots.*

In certain cases the comparison of the sign changes in two sequences allows one to sharpen the estimate of the number of roots as compared with the estimate given by the Fourier-Budan theorem. The first to formulate this type of theorem was Newton but it was proved (by Sylvester) much later, in 1871. Let us replace the sequence $f(x), f'(x), \dots, f^{(n)}(x)$ by the sequence $f_0(x), f_1(x), \dots, f_n(x)$, where

$$f_i(x) = \frac{(n-i)!}{n!} f^{(i)}(x), \tag{2}$$

and consider one more sequence $F_0(x), F_1(x), \dots, F_n(x)$, where $F_0(x) = F(x)$, $F_n(x) = f_n^2(x)$ and

$$F_i(x) = f_i^2(x) - f_{i-1}(x) f_{i+1}(x), \quad i = 1, \dots, n-1. \tag{3}$$

Convention 1.4.1 *Let us take into account only the pairs $f_i(x)$, $f_{i+1}(x)$ for which* $\operatorname{sgn} F_i(x) = \operatorname{sgn} F_{i+1}(x)$.

Let $N_+(x)$ be the number of pairs for which $\operatorname{sgn} f_i(x) = \operatorname{sgn} f_{i+1}(x)$ and let $N_-(x)$ be the number of pairs for which $\operatorname{sgn} f_i(x) = -\operatorname{sgn} f_{i+1}(x)$.

Theorem 1.4.2 (Newton-Sylvester). *Let f be a polynomial of degree n without multiple roots. Then the number of roots of f between a and b, where $a < b$ and $f(a)f(b) \neq 0$, does not exceed either $N_+(b) - N_+(a)$ or $N_-(a) - N_-(b)$.*

Proof. First consider the case when f satisfies the following conditions:

1) no two consecutive polynomials f_i have common roots;
2) no two consecutive polynomials F_i have common roots;
3) the roots of f_i and F_i are distinct from a and b.

In this case formula (3) implies that f_i and F_i have no common roots.

It is easy to derive from (2) and (3) that

$$f_i' = (n-i) f_{i+1}, \tag{4}$$

$$f_i F_i' = (n-i-1)(F_i f_{i+1} + F_{i+1} f_i). \tag{5}$$

Let x move from a to b. The numbers $N_\pm(x)$ only vary if x passes either through a root of f_i or through a root of F_i. Consider separately the following three cases.

Case 1: the passage through a root x_0 of $f_0 = f$. If $f_0(x_0) = 0$, then

$$F_1(x_0) = f_1^2(x_0) - f_0(x_0) f_2(x_0) = f_1^2(x_0) > 0.$$

Therefore the passage through x_0 does not involve a change of sign in the sequence $F_0(x) = 1,\ F_1(x)$. Formula (4) implies that $\operatorname{sgn} f'(x) = \operatorname{sgn} f_1(x)$. Therefore, if $f_1(x_0) > 0$, then $f_0(x_0 - \varepsilon) < 0$ and $f_0(x_0 + \varepsilon) > 0$, whereas if $f_1(x_0) < 0$, then $f_0(x_0 - \varepsilon) > 0$ and $f_0(x_0 + \varepsilon) < 0$. In both cases

$$f_0(x_0 - \varepsilon) f_1(x_0 - \varepsilon) < 0 \quad \text{and} \quad f_0(x_0 + \varepsilon) f_1(x_0 + \varepsilon) > 0.$$

Thus, the passage through x_0 increases N_+ by 1 and decreases N_- by 1. (We only consider the contribution to $N_\pm$ of the pair f_0, f_1.)

Case 2: the passage through a root x_0 of the polynomial f_i, where $i \geq 1$. In this case the change of signs occurs in the sequence f_{i-1}, f_i, f_{i+1}. The possible variants of the signs of the polynomials considered at $x = x_0 \pm \varepsilon$ are considerably restricted by the following relations:

$$\begin{aligned} &1) \operatorname{sgn} f_{i+1} = \operatorname{sgn} f_i' \quad \text{due to (4);} \\ &2) \operatorname{sgn} F_i(x_0) = \operatorname{sgn}\left(f_i^2(x_0) - f_{i-1}(x_0)f_{i+1}(x_0)\right) = \\ &\quad - \operatorname{sgn}\left(f_{i-1}(x_0)f_{i+1}(x_0)\right) \quad \text{due to (3);} \\ &3) \operatorname{sgn} F_{i\pm 1} = \operatorname{sgn} f_{i\pm 1}^2 = 1. \end{aligned}$$

If $F_i(x_0) < 0$ the sign changes occur in the pairs F_{i-1}, F_i and F_i, F_{i+1} but, by Convention 1 just before the theorem, we do not consider such pairs. If $F_i(x_0) > 0$, then $f_{i-1}(x_0)f_{i+1}(x_0) < 0$. The signs of the polynomials f_{i-1}, f_i, f_{i+1} considered at $x = x_0 \pm \varepsilon$ are completely determined by the sign of $f_{i+1}(x_0)$. For both values of the signs, the pairs $f_{i-1}(x_0 - \varepsilon)$, $f_i(x_0 - \varepsilon)$ and $f_i(x_0 - \varepsilon)$, $f_{i+1}(x_0 - \varepsilon)$ contribute to N_+ and N_-, respectively, and then the pairs $f_{i-1}(x_0 + \varepsilon)$, $f_i(x_0 + \varepsilon)$ and $f_i(x_0 + \varepsilon)$, $f_{i+1}(x_0 + \varepsilon)$ contribute the other way round to N_- and N_+, respectively. Thus, their total contribution to N_+ as well as to N_- does not vary.

Case 3: passage through a root x_0 of F_i. In this case the signs of the polynomials satisfy the following relations:

$$\begin{aligned} &1) f_{i-1}(x_0)f_{i+1}(x_0) = f_i^2(x_0) - F_i(x_0) = f_i^2(x_0) > 0; \\ &2) \operatorname{sgn} f_i' = \operatorname{sgn} f_{i+1}; \\ &3) \text{ formula (5) implies that } \operatorname{sgn} F_i' = \operatorname{sgn} f_{i-1}f_{i+1}F_{i+1}. \end{aligned}$$

An easy perusal of the possible scenarios shows that either both N_+ and N_- do not vary, or N_+ increases by 2, or N_- decreases by 2.

It remains to explain how to get rid of conditions 1)–3) imposed on f. If some of these conditions are not satisfied, then after a small variation of the coefficients of f these conditions will be satisfied. But the roots of f are simple ones, and therefore the number of roots of f lying strictly inside the segment $[a, b]$ does not vary under a small variation of the coefficients. □

Remark. For the polynomial f with multiple roots, one should make use of a slightly more subtle argument. Namely, one should consider not arbitrary small variations but only those for which the real root of multiplicity r splits into r distinct real roots. To produce such a small variation, it is more convenient to modify the roots of the polynomial rather than its coefficients.

1.4.2 Sturm's Theorem

Consider the polynomials $f(x)$ and $f_1(x) = f'(x)$. Let us seek the greatest common divisor of f and f_1 with the help of Euclid's algorithm:

$$
\begin{aligned}
f &= q_1 f_1 - f_2,\\
f_1 &= q_2 f_2 - f_3,\\
&\dots\dots\dots\dots\dots\\
f_{n-2} &= q_{n-1} f_{n-1} - f_n,\\
f_{n-1} &= q_n f_n.
\end{aligned}
$$

The sequence $f, f_1, \dots, f_{n-1}, f_n$ is called the *Sturm sequence* of the polynomial f.

Theorem 1.4.3 (Sturm). *Let $w(x)$ be the number of sign changes in the sequence*

$$f(x), \quad f_1(x), \ \dots, \ f_n(x).$$

The number of the roots of f (without taking multiplicities into account) confined between a and b, where $f(a) \neq 0$, $f(b) \neq 0$ and $a < b$, is equal to $w(a) - w(b)$.

Proof. First, consider the case when the roots of f are simple (i.e., the polynomials f and f' have no common roots). In this case f_n is a nonzero constant.

Let us verify first of all that when we pass through one of the roots of polynomials $f_1, \dots, f_{n-1}$ the number of sign changes does not vary. In the case considered, the neighboring polynomials have no common roots, i.e., if $f_r(\alpha) = 0$, then $f_{r\pm1}(\alpha) \neq 0$. Moreover, the equality $f_{r-1} = q_{r-1} f_r - f_{r+1}$ implies that $f_{r-1}(\alpha) = -f_{r+1}(\alpha)$. But in this case the number of sign changes in the sequence $f_{r-1}(\alpha)$, ε, $f_{r+1}(\alpha)$ is equal to 2 both for $\varepsilon > 0$ and for $\varepsilon < 0$.

Let us move from a to b. If we pass through a root x_0 of f, then first the numbers $f(x)$ and $f'(x)$ are of different signs and then they are of the same sign. Therefore the number of sign changes in the Sturm sequence diminishes by 1. All the other sign changes, as we have already shown, are preserved during the passage through x_0.

Now consider the case when x_0 is a root of multiplicity m of f. In this case f and f_1 have a common divisor $(x - x_0)^{m-1}$, and hence the polynomials are divisible by $(x - x_0)^{m-1}$. Having divided f, f_1, $\dots$, f_r by $(x - x_0)^{m-1}$ we obtain the Sturm sequence $\varphi, \varphi_1, \dots, \varphi_r$ for the polynomial $\varphi(x) = \dfrac{f(x)}{(x - x_0)^{m-1}}$. The root x_0 is a simple one for φ, and hence the passage through x_0 increases the number of sign changes in the sequence $\varphi, \varphi_1, \dots, \varphi_r$ by 1. But for a fixed x the sequence f, f_1, $\dots$, f_r is obtained from φ, φ_1, $\dots$, φ_r by multiplication by a constant, and therefore the numbers of sign changes in these sequences coincide. □

1.4.3 Sylvester's theorem

To compute the Sturm sequence is rather a laborious task. Sylvester suggested the following more elegant method for computing the number of the real roots of the polynomial. Let f be a real polynomial of degree n with simple roots $\alpha_1, \dots, \alpha_n$. Set $s_k = \alpha_1^k + \dots + \alpha_n^k$. (Clearly, to calculate s_k one does not have to know the roots of the polynomial because s_k, being a symmetric function, is expressed in terms of the coefficients of the polynomial.)

Theorem 1.4.4 (Sylvester). a) *The number of the real roots of f is equal to the signature of the quadratic form with the matrix*

$$\begin{pmatrix} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \dots & s_{2n} \end{pmatrix}.$$

b) *All the roots of f are positive if and only if the matrix*

$$\begin{pmatrix} s_1 & s_2 & \dots & s_n \\ s_2 & s_3 & \dots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \dots & s_{2n+1} \end{pmatrix}.$$

is positive definite.

Proof. (Hermite) Let ρ be a real parameter. Consider the quadratic form

$$F(x_1, \dots, x_n) = \frac{y_1^2}{\alpha_1 + \rho} + \dots + \frac{y_n^2}{\alpha_n + \rho}, \tag{1.1}$$

$$\text{where } y_r = x_1 + \alpha_r x_2 + \dots + \alpha_r^{n-1} x_n. \tag{1.2}$$

The coefficients of the polynomial F are symmetric functions in the roots of f, and hence they are real. In particular, this means that the form F can be represented as

$$h_1^2 + \dots + h_p^2 - h_{p+1}^2 - \dots - h_n^2,$$

where $h_1, \dots, h_n$ are linear forms in $x_1, \dots, x_n$ with real coefficients.

To the real root α_r there corresponds the summand

$$\frac{y_r^2}{\alpha_r + \rho} = \frac{(x_1 + \alpha_r x_2 + \dots + \alpha_r^{n-1} x_n)^2}{\alpha_r + \rho}.$$

This summand can be represented in the form $\pm h_r^2$, where the plus sign is taken if $\alpha_r + \rho > 0$ and the minus sign otherwise.

The contribution of a pair of conjugate roots α_r and α_s is equal to

$$F_{r,s} = \frac{y_r^2}{\alpha_r + \rho} + \frac{y_s^2}{\alpha_r + \rho}.$$

Let $y_r = u + iv$ and $\dfrac{y}{\alpha_r + \rho} = \lambda + i\mu$, where u, v, λ, μ are real numbers. Then $y_s = u - iv$ and $\dfrac{y}{\alpha_s + \rho} = \lambda - i\mu$. Therefore

$$F_{r,s} = 2\lambda(u^2 - v^2) - 4\mu uv.$$

For $u = 0$ and for $v = 0$, the values of $F_{r,s}$ have opposite signs. Hence after a change of variables we may assume that $F_{r,s} = u_1^2 - v_1^2$.

As a result we see that all the roots of f are real and satisfy the inequality $\alpha_r > -\rho$ if and only if the form (1) is positive definite. The matrix elements of this form are

$$a_{ij} = \frac{\alpha_1^{i+j-2}}{\alpha_1 + \rho} + \cdots + \frac{\alpha_n^{i+j-2}}{\alpha_n + \rho}.$$

Statements a) and b) are obtained by going to the limit as $\rho \longrightarrow +\infty$ and taking $\rho = 0$, respectively. □

The quadratic form that appears in Sylvester's theorem has quite an interesting interpretation. This interpretation will enable us to obtain another proof of Sylvester's theorem; moreover, even for polynomials with multiple roots.

Consider the linear space $V = \mathbb{R}[x]/(f)$ consisting of polynomials considered modulo a polynomial $f \in \mathbb{R}[x]$. We assume that f is monic and $\deg f = n$. The polynomials $1, x, \dots, x^{n-1}$ form a basis of V. To every $a \in V$, we may assign a linear map $V \to V$ given by the formula $v \mapsto av$ (since the elements of V are polynomials we can multiply them). Let $\operatorname{tr}(a)$ be the trace of this map. Consider the symmetric bilinear form

$$\varphi(v, w) = \operatorname{tr}(vw).$$

Theorem 1.4.5. a) *Let $f(x) = (x - \alpha_1) \cdot \dots \cdot (x - \alpha_n) \in \mathbb{R}[x]$ and $s_k = \alpha_1^k + \cdots + \alpha_n^k$. The matrix of φ in the basis $1, x, \dots, x^{n-1}$ has the form*

$$\begin{pmatrix} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \dots & s_{2n} \end{pmatrix}.$$

b) *The signature of the form φ is equal to the number of distinct real roots of f.*

Proof. a) Over $\mathbb{C}$, the polynomial f can be factorized into the product of relatively prime linear factors $f = f_1^{m_1} \cdot \dots \cdot f_r^{m_r}$. Thanks to the Chinese remainder theorem (Lemma on p. 69) the map

$$h \pmod f \mapsto (h \pmod{f_1^{m_1}}, \dots, h \pmod{f_r^{m_r}})$$

determines a canonical isomorphism

$$\mathbb{C}[x]/(f) \cong \mathbb{C}[x]/(f_1^{m_1}) \times \cdots \times \mathbb{C}[x]/(f_r^{m_r}).$$

In this decomposition the factors are orthogonal with respect to φ. Indeed, let polynomials h_i and h_j correspond to factors with distinct numbers i and j, i.e., $h_i \equiv 0 \pmod{f/f_i^{m_i}}$ and $h_j \equiv 0 \pmod{f/f_j^{m_j}}$. Then $h_i h_j \equiv 0 \pmod{f}$, and therefore the map $v \mapsto h_i h_j v$ is the zero one. Hence its trace vanishes. Therefore $\varphi = \varphi_1 + \cdots + \varphi_r$, where φ_i is the restriction of φ onto the subspace $\mathbb{C}[x]/(f_i^{m_i}) = \mathbb{C}[x]/(x-\alpha_i)^{m_i}$. It remains to verify that $\varphi_i(1, x^k) = m_i \alpha_i^k$.

It is easy to calculate the matrix of the form φ_i in the basis

$$1,\ x-\alpha_i,\ \ldots,\ (x-\alpha_i)^{m_i-1}.$$

Indeed, in this basis the map $v \mapsto (x-\alpha_i)^k v$ has a triangular matrix; and the trace of this matrix is equal to m_i if $k = 0$ and to 0 if $k > 0$. Since

$$0 = \varphi_i(1, x-\alpha_i) = \varphi_i(1, x) - \alpha_i \varphi_i(1, 1) = \varphi_i(1, x) - m_i \alpha_i,$$

it follows that $\varphi_i(1, x) = m_i \alpha_i$. Next, with the help of the equality

$$\varphi_i\left(1,\ (x-\alpha_i)^k\right) = 0$$

and induction on k we see that $\varphi_i(1, x^k) = m_i \alpha_i^k$.

b) Computing the signature we must remain in $\mathbb{R}$, and therefore we decompose f over $\mathbb{R}$ into the product of relatevely prime linear or quadratic factors: $f = f_1^{m_1} \cdot \ldots \cdot f_r^{m_r}$. Again consider the decomposition

$$\mathbb{R}[x]/(f) \cong \mathbb{R}[x]/(f_1^{m_1}) \times \cdots \times \mathbb{R}[x]/(f_r^{m_r}).$$

It suffices to verify that the signature of the restriction of φ onto $\mathbb{R}[x]/(f_i^{m_i})$ is equal to 1 if $\deg f_i = 1$ and to 0 if f_i is an irreducible over $\mathbb{R}$ polynomial of degree 2. As we have already established, in the basis $1,\ x-\alpha_i,\ (x-\alpha_i)^{m_i-1}$, the matrix of φ_i is equal to

$$\begin{pmatrix} m_i & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Therefore if $\deg f_i = 1$ the signature of φ_i is equal to 1.

If f_i is an irreducible over $\mathbb{R}$ polynomial of degree 2, then $\mathbb{R}[x]/(f_i^{m_i}) \cong \mathbb{R}[x]/(x^2+1)^{m_i}$. Here we mean an isomorphism over $\mathbb{R}$. Therefore it suffices to calculate the signature of φ on $\mathbb{R}[x]/(x^2+1)^m$. It is convenient to calculate the matrix of φ in the basis

$$1,\ x^2,\ x^2+1,\ x(x^2+1),\ (x^2+1)^2, \ldots, x(x^2+1)^{m-1},\ (x^2+1)^{m-1}.$$

In this basis, the operators of multiplication by x and x^2 have matrices

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ -1 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & -1 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 & 0 & 1 & 0 & 0 & \dots \\ 0 & -1 & 0 & 1 & 0 & \dots \\ 0 & 0 & -1 & 0 & 1 & \dots \\ 0 & 0 & 0 & -1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

respectively. Therefore the trace of the operator of multiplication by x is equal to 0 and the trace of the operator of multiplication by x^2 is equal to $-2m$. The operators of multiplication by $x^a(x^2+1)^k$, where $a = 0,\ 1,\ 2$ and $k \geq 1$, are represented by diagonal matrices with zero main diagonals; their traces vanish. As a result, we see that the matrix of the form φ is equal to

$$\begin{pmatrix} 2m & 0 & 0 & \dots & 0 \\ 0 & -2m & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

The signature of such a form is equal to zero. □

1.4.4 Separation of complex roots

Sturm's theorem enables one to indicate algorithmically a set of segments that contain all the real roots of a real polynomial and, moreover, each such segment contains precisely one root. In a series of papers (1869–1878), Kronecker developed a theory with an algorithm to indicate a set of disks which contain all the complex roots of a complex polynomial so that each disk contains exactly one root. More exactly, Kronecker showed that the number of complex roots inside the given disk can be computed with the help of Sturm's theorem.

Let $z = x + iy$. Let us represent the polynomial $P(z)$ in the form $P(z) = \varphi(x, y) + i\psi(x, y)$. We will assume that P has no multiple roots, i.e., if $P(z) = 0$, then $P'(z) \neq 0$.

To every root of P, there corresponds the intersection point of the curves $\varphi = 0$ and $\psi = 0$. Therefore the number of roots of P lying inside a closed non-self-intersecting curve γ is equal to the number of the intersection points of the curves $\varphi = 0$ and $\psi = 0$ lying inside γ. This number can be calculated as follows. Let us circumscribe the curve γ in the positive direction, i.e., counterclockwise, and to each intersection point of the curves γ and $\varphi = 0$ we assign the number $\varepsilon_i = \pm 1$ according to the following rule: $\varepsilon_i = 1$ if we move from the domain $\varphi\psi > 0$ to the domain $\varphi\psi < 0$, or $\varepsilon_i = -1$ if, the other way round, we move from the domain $\varphi\psi < 0$ to the domain $\varphi\psi > 0$.

In the general position the number of intersection points of the curves γ and $\varphi = 0$ is even (since at every intersection point the function φ changes its sign), and hence $\sum \varepsilon_i = 2k$, where k is an integer.

Theorem 1.4.6 (Kronecker). a) *The number k is equal to the number of intersection points of the curves $\varphi = 0$ and $\psi = 0$ lying inside the curve γ.*

b) *If γ is a circle of given radius with given center, then for the given polynomial P the number k can be algorithmically computed.*

Proof. a) Clearly, $dP(z) = (\varphi_x + i\psi_x)dx + (\psi_y - i\varphi_y)i\,dy$. Hence

$$\varphi_x + i\psi_x = P'(z) = \psi_y - i\varphi_y,$$

and therefore

$$\psi_y = \varphi_x \text{ and } \psi_x = -\varphi_y$$

(the Cauchy-Riemann relations). Therefore

$$\begin{vmatrix} \phi_x & \phi_y \\ \psi_x & \psi_y \end{vmatrix} = \begin{vmatrix} \phi_x & \phi_y \\ \phi_y & -\phi_x \end{vmatrix} = \phi_x^2 + \varphi_y^2 > 0.$$

This means that the rotation from the vector $\operatorname{grad}\varphi = (\varphi_x, \varphi_y)$ to the vector $\operatorname{grad}\psi = (\psi_x, \psi_y)$ is a counterclockwise one. Geometrically this means that the domains $\varphi\psi > 0$ and $\varphi\psi < 0$ are positioned as shown in Fig. 1.1.

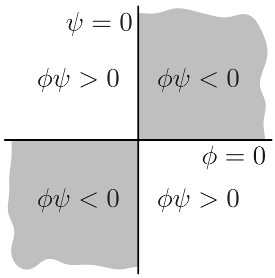

FIGURE 1.1

Let us contract the curve γ into a point. Under the passage through the intersection point of the curves $\varphi = 0$ and $\psi = 0$ the number k diminishes by 1 (Fig. 1.2) and under the reconstruction depicted on Fig. 1.3 the number k does not vary. It is also clear that when the curve becomes sufficiently small it does not intersect the curves $\varphi = 0$ and $\psi = 0$, and in this case $k = 0$.

b) The circle of radius r and center (a, b) can be parameterized with the real parameter t as follows:

$$x = a + r\frac{1 - t^2}{1 + t^2}, \quad y = b + r\frac{2t}{1 + t^2}.$$

Having substituted these expressions into $\varphi(x, y)$ we obtain a polynomial $\Phi(t)$ with real coefficients. The real roots of this polynomial correspond to the

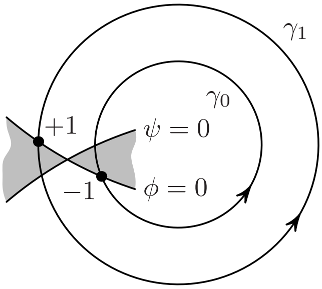

FIGURE 1.2

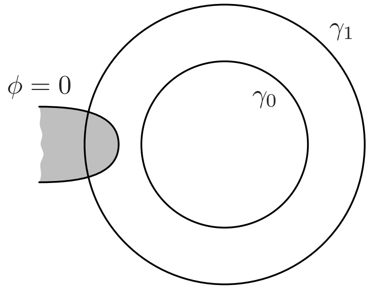

FIGURE 1.3

intersection points of the curves γ and $\varphi = 0$. By Sturm's theorem, for every root, we can find a segment that contains it. Having calculated the sign of the function $\varphi\psi$ at the endpoints of this segment one can find the corresponding numbers ε_i. □

1.5 Lagrange's series and estimates of the roots of a given polynomial

1.5.1 The Lagrange-Bürmann series

Recall that if $f(z) = \sum\limits_{n=-\infty}^{\infty} c_n(z-a)^n$, then

$$\frac{1}{2\pi i}\int\limits_{\gamma} f(z)\,dz = c_{-1},$$

where γ is any curve circumscribing point a. We will use this fact to obtain the expansion of the function $f(z)$ into a series in powers of $\varphi(z) - b$, where $b = \varphi(a)$. To be able to do so, the function $\varphi(z)$ should be invertible in a neighborhood of a, i.e., $\varphi'(a) \neq 0$. If $\varphi(z)$ is invertible, then

$$\frac{f'(z)\varphi'(a)}{\varphi(z)-\varphi(a)} = \frac{f'(z)\varphi'(a)}{\varphi'(a)(z-a)+\cdots} = \frac{f'(a)}{z-a} + \ldots,$$

and hence

$$\frac{1}{2\pi i}\int\limits_{\gamma} \frac{f'(z)\varphi'(a)}{\varphi(z)-\varphi(a)}\,dz = f'(a).$$

Having integrated this identity, we obtain

$$f(z) - f(a) = \int_a^z f'(\zeta)\,d\zeta = \frac{1}{2\pi i} \int_a^z \int_\gamma \frac{f'(w)\varphi'(\zeta)}{\varphi(w) - \varphi(\zeta)}\,dw\,d\zeta.$$

Let us transform the expression obtained having separated the terms $\varphi(z) - b$, where $b = \varphi(a)$:

$$\frac{f'(w)\varphi'(\zeta)}{\varphi(w) - \varphi(\zeta)} = \frac{f'(w)\varphi'(\zeta)}{\varphi(w) - b} \cdot \frac{\varphi(w) - b}{\varphi(w) - \varphi(\zeta)},$$

$$\frac{\varphi(w) - b}{\varphi(w) - \varphi(\zeta)} = \left(1 - \frac{\varphi(\zeta) - b}{\varphi(w) - b}\right)^{-1} = \sum_{m=0}^{\infty} \left(\frac{\varphi(\zeta) - b}{\varphi(w) - b}\right)^m.$$

By changing the order of integration we obtain

$$f(z) - f(a) = \frac{1}{2\pi i} \int_\gamma \left(\int_a^z \frac{f'(w)\varphi'(\zeta)}{\varphi(w) - b} \sum_{m=0}^{\infty} \left(\frac{\varphi(\zeta) - b}{\varphi(w) - b}\right)^m d\zeta \right) dw.$$

When we calculate the integral over ζ we only need the factors depending on ζ:

$$\int_a^z \varphi'(z)\,(\varphi(\zeta) - b)^m\,d\zeta = \int_{\varphi(a)}^{\varphi(z)} (\varphi(\zeta) - b)^m\,d\varphi(\zeta) = \frac{(\varphi(\zeta) - b)^{m+1}}{m+1}$$

(we have taken into account that $\varphi(a) - b = 0$).

Thus,

$$f(z) - f(a) = \sum_{m=0}^{\infty} \frac{(\varphi(\zeta) - b)^{m+1}}{m+1} \frac{1}{2\pi i} \int_\gamma \frac{f'(w)\,dw}{(\varphi(w) - b)^{m+1}}.$$

Consider a function $\psi(w)$ such that $\dfrac{1}{\varphi(w) - b} = \dfrac{\psi(w)}{w - a}$, i.e.,

$$\psi(w) = \frac{w - a}{\varphi(w) - b}. \tag{1}$$

For this function $\psi(w)$, we have

$$\frac{1}{2\pi i} \int_\gamma \frac{f'(w)\,dw}{(\varphi(w) - b)^{m+1}} = \frac{1}{2\pi i} \int_\gamma \frac{f'(w)\,(\psi(w))^{m+1}\,dw}{(w - a)^{m+1}} =$$

$$= \frac{1}{m!} \cdot \frac{d^m}{dw^m} \Big(f'(w)\,(\psi(w))^{m+1} \Big)_{w=a}.$$

Indeed,

$$f'(w)\,(\psi(w))^{m+1} = \sum_{k=0}^{\infty} c_k (w-a)^k,$$

where

$$c_k = \frac{1}{k!}\cdot\frac{d^k}{dw^k}\Big(f'(w)\,(\psi(w))^{k+1}\Big)_{w=a}.$$

The integral we are interested in is equal to c_m — the coefficient of $(w-a)^{-1}$ in the series $\sum\limits_{k=0}^{\infty} c_k(w-a)^{k-m-1}$.

As a result, we obtain the following expansion of $f(z)$ into powers of $\varphi(z)-b$:

$$f(z)-f(a) = \sum_{n=1}^{\infty} \frac{(\varphi(z)-b)^n}{n!}\cdot\frac{d^{n-1}}{dw^{n-1}}\Big(f'(w)\,(\psi(w))^n\Big)_{w=a}, \qquad (2)$$

where $\psi(w)$ is given by formula (1). The series (2) is called *Bürmann's series.* Bürmann obtained it in 1799 while generalizing a series Lagrange obtained in 1770. The *Lagrange series* can be obtained from Bürmann's series for $\varphi(z) = \dfrac{z-a}{h(z)}$, where $h(z)$ is a function. In this case $b=\varphi(a)=0$ and

$$\psi(z) = \frac{z-a}{\varphi(z)-b} = h(z).$$

Therefore

$$f(z) = f(a) + \sum_{n=0}^{\infty}\frac{s^n}{n!}\cdot\frac{d^{n-1}}{da^{n-1}}\Big(f'(a)\,(h(a))^n\Big),$$

where $s=\varphi(z)$. In particular,

$$z = a + \sum_{n=0}^{\infty}\frac{s^n}{n!}\cdot\frac{d^{n-1}}{da^{n-1}}\,(h(a))^n. \qquad (3)$$

Thus, if the series (3) converges, it enables one to calculate the roots of the equation

$$z = a + s\,h(z).$$

Example. Let $h(z)=\dfrac{1}{z}$. In this case the series (3) has the form

$$z = a + \sum_{n=1}^{\infty}\frac{(-1)^{n-1}(2n-2)!}{n!\,(n-1)!\,a^{2n-1}}\,s^n. \qquad (4)$$

Series (4) converges for $|s| < \dfrac{|a|^2}{4}$. The equation under consideration,

$$z = a + \frac{s}{z},$$

has two roots

$$\frac{a}{2}\left(1+\sqrt{1+\frac{4s^2}{a^2}}\right) \text{ and } \frac{a}{2}\left(1-\sqrt{1+\frac{4s^2}{a^2}}\right).$$

The series (4) represents only the first of these roots.

1.5.2 Lagrange's series and estimation of roots

Lagrange's series enables one in certain cases to estimate the roots of polynomials. Consider, for example, the polynomial

$$f(z) = a_0 + a_1(z-c) + a_2(z-c)^2 + \cdots + a_k(z-c)^k.$$

The equation $f(z) = 0$ can be expressed in the form

$$z = c + s\ h(z),$$

where $s = -\dfrac{1}{a_1}$ and $h(z) = a_0 + a_2(z-c)^2 + a_3(z-c)^3 + \cdots + a_k(z-c)^k$. Lagrange's series for this equation is of the form

$$z = c + \sum_{n=1}^{\infty} \frac{s^n}{n!} \cdot \frac{d^{n-1}}{dz^{n-1}}\ \left(h^n(z)\right)_{z=c}.$$

In our case

$$h^n(z) = \sum_{\nu_0+\nu_2+\cdots+\nu_k=n} a_0^{\nu_0} a_2^{\nu_2} \cdot \ldots \cdot a_k^{\nu_k} \frac{n!}{\nu_0!\, \nu_2! \cdot \ldots \cdot \nu_k!} (z-c)^{2\nu_2+\cdots+k\nu_k},$$

and hence

$$\frac{d^{n-1}}{dz^{n-1}}\ \left(h^n(z)\right)_{z=c} = \sum \frac{(n-1)!}{\nu_0!\, \nu_2! \cdot \ldots \cdot \nu_k!} a_0^{\nu_0} a_2^{\nu_2} \cdot \ldots \cdot a_k^{\nu_k}, \tag{1}$$

where the sum runs over the collections $\{\nu_0, \nu_2, \ldots, \nu_k\}$ such that

$$\nu_0 + \nu_2 + \cdots + \nu_k = n, \quad 2\nu_2 + \cdots + k\nu_k = n-1.$$

These relations are equivalent to the relations

$$n - 1 = 2\nu_2 + \cdots + k\nu_k, \quad \nu_0 = \nu_2 + 2\nu_3 + \cdots + (k-1)\nu_k + 1.$$

Since $s = -\dfrac{1}{a_1}$, we obtain

$$z = c - \frac{a_0}{a_1} \sum \frac{(2\nu_2 + \cdots + k\nu_k)!}{\nu_0!\, \nu_2! \cdot \ldots \cdot \nu_k!} \left(\frac{a_0 a_2}{(-a_1)^2}\right)^{\nu_2} \cdot \ldots \cdot \left(\frac{a_0^{k-1} a_k}{(-a_1)^k}\right)^{\nu_k}, \tag{2}$$

where $\nu_0 = \nu_2 + 2\nu_3 + \cdots + (k-1)\nu_k + 1$.

If the series (2) converges, the number z so determined is one of the roots of the equation $f(z) = 0$.

Theorem 1.5.1 ([Be3]). *Let $|a_0| + |a_2| + \cdots + |a_k| < |a_1|$. Then the series* (2) *converges and the root z determined by the series satisfies*

$$|z - c| \leq -\ln\left(1 - \frac{1}{|a_1|}\left(|a_0| + |a_2| + \cdots + |a_k|\right)\right).$$

Proof. Formula (1) implies that

$$\left|\frac{1}{n!} \cdot \frac{d^{n-1}}{dz^{n-1}}\left(h^n(z)\right)_{z=c}\right| \leq \frac{1}{n}\left(|a_0| + |a_2| + \cdots + |a_k|\right)^n.$$

Hence

$$|z - c| \leq \sum_{n=1}^{\infty} \frac{|a_1|^{-n}}{n}\left(|a_0| + |a_2| + \cdots + |a_k|\right)^n =$$
$$-\ln\left(1 - \frac{1}{|a_1|}\left(|a_0| + |a_2| + \cdots + |a_k|\right)\right). \square$$

1.6 Problems to Chapter 1

1.1 Prove that a polynomial $f(x)$ is divisible by $f'(x)$ if and only if $f(x) = a_0(x - x_0)^n$.

1.2 Prove that the polynomial

$$a_0 + a_1 x^{m_1} + a_2 x^{m_2} + \cdots + a_n x^{m_n}$$

has at most n positive roots.

1.3 [Newton] Prove that if all the roots of the polynomial

$$P(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n$$

with real coefficients are real and distinct, then

$$a_i^2 > \frac{n-i+1}{n-1} \cdot \frac{i+1}{i} a_{i-1} a_{i+1} \quad \text{for } i = 1, 2, \ldots, n-1.$$

1.4 Prove that the polynomial

$$a_1 x^{m_1} + a_2 x^{m_2} + \cdots + a_n x^{m_n}$$

has no nonzero roots of multiplicity greater than $n - 1$.

1.5 Find the number of real roots of the following polynomials:

a) $1 + x + \dfrac{x^2}{2} + \cdots + \dfrac{x^n}{n}$;

b) $nx^n - x^{n-1} - \cdots - 1$.

1.6 Let $0 = m_0 < m_1 < \cdots < m_n$ and $m_i \equiv i \pmod 2$. Prove that the polynomial

$$a_0 + a_1x^{m_1} + a_2x^{m_2} + \cdots + a_nx^{m_n}$$

has at most n real roots.

1.7 Let x_0 be a root of the polynomial $x^n + a_1x^{n-1} + \cdots + a_n$. Prove that for any $\varepsilon > 0$ there exists a $\delta > 0$ such that if $|a_i - a_i'| < \delta$ for $i = 1, \dots, n$, then the polynomial $x^n + a_1'x^{n-1} + \cdots + a_n'$ has a root x_0' such that $|x_0 - x_0'| < \varepsilon$.

1.8 Let the numbers $a_1, \dots, a_n$ be distinct and let the numbers $b_1, \dots, b_n$ be positive. Prove that all the roots of the equation

$$\sum \frac{b_k}{x - a_k} = x - c, \quad \text{where } \ c \in \mathbb{R},$$

are real.

1.9 Find all the roots of the equation

$$\frac{(x^2 - x + 1)^3}{x^2(x-1)^2} = \frac{(a^2 - a + 1)^3}{a^2(a-1)^2}.$$

1.10 Find the number of roots of the polynomial $x^n + x^m - 1$ whose absolute values are less than 1.

1.11 Let $f(z) = z^n + a_1z^{n-1} + \cdots + a_n$, where $a_1, \dots, a_n \in \mathbb{C}$. Prove that any root z of f satisfies $-\beta \le \operatorname{Re} z \le \alpha$, where α is the only positive root of the polynomial

$$x^n + (\operatorname{Re} a_1)x^{n-1} - |a_2|x^{n-2} - \cdots - |a_n|$$

and β is the only positive root of the polynomial

$$x^n - (\operatorname{Re} a_1)x^{n-1} - |a_2|x^{n-2} - \cdots - |a_n|.$$

1.12 [Su1] Let $f(z)$ be a polynomial of degree n with complex coefficients. Prove that the polynomial $F = f \cdot f' \cdot f'' \cdot \ldots \cdot f^{(n-1)}$ has at least $n+1$ distinct roots.

1.7 Solutions of selected problems

1.3. Set $Q(y) = y^nP(y^{-1})$. The roots of $Q(y)$ are also real and distinct. Hence the roots of the quadratic polynomial

$$Q^{(n-2)}(y) = (n-2)\cdot(n-3)\cdot\dots\cdot 4\cdot 3\ \big(n(n-1)a_ny^2 + 2(n-1)a_{n-1}y + 2a_{n-2}\big)$$

are real and distinct. Therefore

$$(n-1)^2 a_{n-1}^2 > 2n(n-1)a_n a_{n-2}.$$

If $i = n-1$ the desired inequality is proved.

Now consider the polynomial

$$P^{(n-i-1)}(x) = b_0 x^{i+1} + b_1 x^i + \dots + b_{i+1}x^2 + b_i x + b_{i-1}.$$

Applying to it the inequality already proved we obtain

$$b_i^2 > \frac{2(i+1)}{i} b_{i-1}b_{i+1}.$$

Since

$$\begin{aligned} b_{i+1} &= (n-i+1)\cdot \ldots \cdot 4 \cdot 3\, a_{i+1}, \\ b_i &= (n-i)\cdot \ldots \cdot 3 \cdot 2\, a_i, \\ b_{i-1} &= (n-i-1)\cdot \ldots \cdot 2 \cdot 1\, a_{i-1}, \end{aligned}$$

it follows that

$$(2(n-i)a_i)^2 > \frac{2(i+1)}{i} 2(n-i+1)(n-i)a_{i-1}a_{i+1}.$$

After simplification we obtain the desired inequality.

1.11. As x grows from 0 to $+\infty$, the function $x^n \pm \operatorname{Re} a_1$ monotonically increases, whereas the function

$$\frac{|a_2|}{x} + \frac{|a_3|}{x^2} + \dots + \frac{|a_n|}{x^{n-1}}$$

monotonically decreases. Therefore each of the polynomials considered has only one positive root.

Let $f(z) = 0$ and $\operatorname{Re} z > \alpha$. Then

$$\begin{aligned} \alpha + \operatorname{Re} a_1 < \operatorname{Re}(z + a_1) \le |z + a_1| = \left| \frac{a_2}{z} + \frac{a_3}{z^2} + \dots + \frac{a_n}{z^{n-1}} \right| \le \\ \le \frac{|a_2|}{|z|} + \dots + \frac{|a_n|}{|z|^{n-1}} < \frac{|a_2|}{\alpha} + \dots + \frac{|a_n|}{\alpha^{n-1}} \end{aligned}$$

(the last inequality follows since $|z| \ge \operatorname{Re} z > \alpha$). On the other hand, by the hypothesis

$$\alpha + \operatorname{Re} a_1 = \frac{|a_2|}{\alpha} + \dots + \frac{|a_n|}{\alpha^{n-1}}:$$

a contradiction.

The estimate of $\operatorname{Re} z$ from below is obtained as the estimate from above of the real part of the root z of $(-1)^n f(-z)$.

1.12. Let $z_1, \dots, z_m$ be the distinct roots of F, and let $\mu_j(r)$ be the multiplicity of z_j as a root of $f^{(r)}$, where $r = 0, 1, \dots, n-1$. Consider the symmetric functions

$$s_k(r) = \sum_{j=1}^{k} \mu_j(r) z_j^k, \tag{1}$$

i.e., $s_k(r)$ is the sum of the kth powers of the roots of $f^{(r)}$. The elementary symmetric functions in the roots of $f^{(r)}$ will be denoted by $\sigma_k(r)$ (for $k > n-r$ we set $\sigma_k(r) = 0$).

It is easy to verify that, if $f(z) = \sum\limits_{k=0}^{n} (-1)^k a_k z^{n-k}$, then

$$f^{(r)}(z) = \sum_{k=0}^{n-r} (-1)^k a_k \frac{(n-k)!}{(n-k-r)!} z^{n-k-r}.$$

Hence

$$\sigma_k(r) = \frac{a_k}{a_0} \cdot \frac{(n-k)!(n-r)!}{n!(n-k-r)!} = \frac{a_k}{a_0} \cdot \frac{(n-k)!}{n!} \cdot (n-r) \cdot \dots \cdot (n-k-r+1).$$

Therefore $\sigma_k(r)$ is a polynomial of degree k in r and $\sigma_k(n) = 0$.

On p. 79, for $k \geq 1$, the identity

$$s_k = \begin{vmatrix} \sigma_1 & 1 & 0 & \dots & 0 \\ 2\sigma_2 & \sigma_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k\sigma_k & \sigma_{k-1} & \sigma_{k-2} & \dots & \sigma_1 \end{vmatrix}$$

is proved. This identity implies, in particular, that $s_k(r)$, where $k > 0$, can be represented as a linear combination of expressions $\sigma_{k_1}(r) \cdots \sigma_{k_p}(r)$, where $k_1 + \dots + k_p = k$, and the coefficients of this linear combination do not depend on r. Therefore, if $k \geq 1$, then $s_k(r)$ is a polynomial in r of degree not greater than k. It is also clear that $s_0(r) = \sum\limits_{j=1}^{m} \mu_j(r) = n - r$ and $s_k(n) = 0$ for all $k \geq 0$.

Consider the relation (1) for $k = 0, 1, \dots, m-1$ as a system of linear equations for unknowns $\mu_j(r)$, where $j = 1, \dots, m$. By the hypothesis, the numbers $z_1, \dots, z_m$ are distinct, and therefore the determinant of the system considered does not vanish (this determinant is a it Vandermond determinant, see [Pr1]). Having solved this system of linear equations via Cramer's algorithm we obtain a representation of $\mu_j(r)$ in the form of a linear combination of the $s_k(r)$, where $k = 0, \dots, m-1$, with coefficients independent of r. Hence $\mu_j(r)$ is a polynomial in r of degree $d_j \leq m - 1$. Since $s_k(n) = 0$ for all k, we have $\mu_j(n) = 0$.

Let the number of distinct roots of F be strictly less than $n + 1$, i.e., $m < n + 1$. Then $d_j \leq m - 1 < n$, i.e., $\mu_j(r)$ is a polynomial in r of degree $\leq n - 1$. In this case

$$\deg \Delta^1 \mu_j(r) = \mu_j(r+1) - \mu_j(r) \le n-2,$$
$$\deg \Delta^2 \mu_j(r) = \Delta^1 \mu_j(r+1) - \Delta^1 \mu_j(r) \le n-3, \ \dots ,$$

$\Delta^{n-1}\mu_j(r)$ is a constant, and $\Delta^n \mu_j(r)$ is identically zero. In particular,

$$\Delta^n \mu_j(0) = \sum_{r=0}^{n} (-1)^r \binom{n}{r} \mu_j(r) = 0.$$

To arrive at a contradiction, it suffices to show that $\Delta^n \mu_1(0) \neq 0$.

Consider the convex hull of the roots of f. By the Gauss-Lucas theorem (Theorem 1.2.1 on p. 13), this convex hull coincides with the convex hull of the points $z_1, \dots, z_m$. We may assume that z_1 is a vertex of the convex hull of the roots of f. Then z_1 lies outside the convex hull of the points $z_2, \dots, z_m$. Let $\mu = \mu_1(0)$ be the multiplicity of z_1 as of a root of f. Then for $0 \le r \le \mu - 1$ the number z_1 is a root of multiplicity $\mu - r$ of $f^{(r)}$ and $f^{(\mu)}(z_1) \neq 0$. The convex hull of the roots of $f^{(\mu)}$ does not contain z_1, and hence $f^{(r)}(z_1) \neq 0$ for $r \ge \mu$. Therefore

$$\mu_1(r) = \begin{cases} \mu - r & \text{for } 0 \le r \le \mu - 1; \\ 0 & \text{for } r \ge \mu. \end{cases}$$

It is also clear that $\mu \le n - 1$, since f has at least one root distinct from z_1. Hence

$$\Delta^2 \mu_1(r) = \begin{cases} 0 & \text{for } 0 \le r \le n-1,\ r \neq \mu - 1; \\ 1 & \text{for } r = \mu - 1. \end{cases}$$

Therefore, for $n > 2$, we obtain

$$\Delta^n \mu_1(0) = \Delta^{n-2} \left(\Delta^2 \mu_1 \right)(0) = \sum_{r=0}^{n-2} (-1)^r \binom{n-2}{r} \Delta^2 \mu_1(r) = (-1)^{\mu-1} \binom{n-2}{\mu-1},$$

and, for $n = 2$, we obtain $\mu = 1$ and $\Delta^2 \mu_1(0) = 1$. In both cases $\Delta^n \mu_1(0) \neq 0$, as was required.

http://www.springer.com/978-3-642-03979-9

Polynomials
Prasolov, V.V.
2004, XIII, 301 p. 8 illus., Softcover
ISBN: 978-3-642-03979-9