Annals of Mathematics
Vol. 51, No. 1, January, 1950

ON THE DISTRIBUTION OF ROOTS OF POLYNOMIALS

By P. Erdös and P. Turán

(Received July 12, 1948)

1. We start by explaining two groups of theorems and we shall derive both from a common source.

P. Bloch and G. Pólya[1] investigated first the question of giving an upper estimation of the number R of real roots of

$$f(x) = a_0 + a_1x + \cdots + a_nx^n \tag{1.1}$$

whenever

$$|a_0| \geqq \mu', \qquad |a_n| \geqq \mu', \qquad |a_\nu| \leqq \mu, \qquad \nu = 1, 2, \cdots, (n-1).$$

They proved that the number of real roots is[2]

$$< A_1(\mu, \mu') \frac{n \log \log n}{\log n}.$$

A few years later Erhardt Schmidt[3] proved the sharper inequality

$$R^2 \leqq A_2(\mu, \mu')n \log \frac{\mu n}{\mu'}$$

and the still sharper one

$$R^2 \leqq A_3\, n \log \frac{|a_0| + \cdots + |a_n|}{\sqrt{|a_0 a_n|}} \equiv A_3\, n \log P, \tag{1.2}$$

where as throughout the present paper,

$$P \equiv \frac{|a_0| + \cdots + |a_n|}{\sqrt{|a_0 a_n|}}. \tag{1.3}$$

His detailed proof has never been published because I. Schur[4] found shortly thereafter an elementary proof for it; his method furnishing at the same time the proof of the inequality

$$R^2 \leqq 4n \log P \tag{1.4}$$

with the best possible constant 4 and that of

$$R^2 - 2R \leqq 2n \log \frac{Q}{2}$$
$$Q = \frac{|a_0|^2 + \cdots + |a_n|^2}{|a_0 a_n|}. \tag{1.5}$$

[1] Proc. Lond. Math. Soc. (2) 33 (1931), 102–114.

[2] We explicitly designate the parameters on which the quantity A_1 and, later, A_2, $A_2 \cdots$ depend. If no dependence is indicated then the corresponding quantity is a numerical constant.

[3] Preuss. Akad. Wiss. Sitzungsber. (1932) 321.

[4] Ibid. (1933), 403–428.

Further G. Szegö[5] found refinements of (1.4), simplifications in the proof of (1.5), and discovered that Schur's extremal-polynomials are essentially Jacobi-polynomials. A further very simple proof for (1.2) was given by Littlewood and Offord[6].

2. R. Jentzsch[7] first proved that if

$$g(z) = 1 + b_1 z + \cdots + b_n z^n + \cdots \tag{2.1}$$

has the the unit-circle as circle of convergence then every point of this circle is a cluster-point of zeros of the partial sums

$$s_n(z) = 1 + b_1 z + \cdots + b_n z^n.$$

G. Szegö[8] found the interesting generalization of this theorem according to which there is an infinite sequence of indices

$$n_1 < n_2 < \cdots < n_k < \cdots$$

such that the roots of $s_{n_k}(z)$ in a fixed annulus

$$1 - \epsilon \leqq |z| \leqq 1 + \epsilon, \qquad \epsilon > 0 \text{ and arbitrarily small}$$

are uniformly dense in the sense of Weyl. This means, as is well known, that denoting the roots of $s_{n_k}(z)$ by

$$z_\nu^{(k)} = r_\nu^{(k)} e^{i\varphi_\nu^{(k)}}, \qquad \nu = 1, 2, \cdots, n_k$$

and given α and β with

$$0 \leqq \alpha < \beta < 2\pi$$

we have

$$\lim_{k\to\infty} \frac{1}{n_k} \sum_{\alpha \leqq \varphi_\nu^{(k)} \leqq \beta} 1 = \frac{\beta - \alpha}{2\pi}. \tag{2.2}$$

3. Now we are going to show how the theorems (1.2) and (2.2) which seem unrelated at the first glance, can be derived from a common source. We state the following

THEOREM I. *If the roots of the polynomial*

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n \tag{3.1}$$

are denoted by

$$z_\nu = r_\nu e^{i\varphi_\nu}, \qquad \nu = 1, 2, \cdots, n \tag{3.2}$$

then for every $0 \leqq \alpha < \beta \leqq 2\pi$ *we have*

$$\left| \sum_{\alpha \leqq \varphi_\nu \leqq \beta} 1 - \frac{\beta - \alpha}{2\pi} n \right| < 16 \sqrt{n \log \frac{|a_0| + \cdots + |a_n|}{\sqrt{|a_0 a_n|}}} = 16\sqrt{n \log P}. \tag{3.3}$$

---

[5] Ibid. (1934), 86–98.

[6] Proc. Cambridge Philos. Soc. 35 (1939), 133–148.

[7] Untersuchungen zur Theorie der Folgen analytischer Functionen. Inaug.-diss., Berlin, 1914.

[8] Berlin Math. Ges. Sitzungsber. 21 (1922), 59–64.

The content of this theorem can be expressed by saying that the roots of a polynomial are uniformly distributed in the different angles with vertex at the origin if the coefficients "in the middle" are not too large compared with the extreme ones. In the case $a_0 = a_1 = \cdots = a_n$ the uniform distribution is of course much more perfect than is expressed in our theorem and represents the ideal case; but our theorem shows that if all coefficients satisfy the condition

$$n^{-\lambda} \leqq |a_\nu| \leqq n^{\lambda} \qquad \nu = 0, 1, \cdots, n \tag{3.4}$$

then

$$P \leqq (n+1)n^{2\lambda} < (n+1)^{2\lambda+1},$$

i.e.

$$\left| \sum_{\alpha \leqq \varphi_\nu \leqq \beta} 1 - \frac{\beta - \alpha}{2\pi} n \right| < 16\sqrt{2\lambda + 1}\sqrt{n \log (n+1)}. \tag{3.5}$$

Hence a rather radical change of the coefficients restricted only by (3.4) cannot "spoil" the uniformly dense distribution of the roots in angles very much.

4. The idea of deducing the theorems of Erhardt-Schmidt and Jentzsch-Szegö from a common source does not seem to be new. I. Schur delivered a lecture in the physical-mathematical section of the Prussian Academy on March3, 1934 as a continuation of his paper[4]. The content of his lecture we know only from the following report:[9] "Es wird gezeigt dass der Satz von Robert Jentzsch über die Nullstellen der Abschnitte einer Potenzreihe mit endlichem Konvergenzradius aus einem allgemeinerem Satz folgt, der ohne Benutzung funktionentheoretischer Hilfsmittel bewiesen werden kann. Die zum Beweise erforderlichen Abschätzungen der Wurzeln einer algebraischen Gleichung werden mit Hilfe des Matrizenkalküls abgeleitet." The fact that the indicated more general theorem contains also the theorem of Erhardt-Schmidt we suspect only from the fact that the lecture was a continuation of his investigations in the paper[4]; nothing has been published about this more general theorem as far as we could find in the Zentralblatt and Math. Reviews. Since our method does not use the matrix-calculus, the method seems to be anyway different from that used by I. Schur.

5. First we deduce from Theorem I Erhardt-Schmidt's inequality (1.2), even in a slightly sharpened form. We apply Theorem I three times, to the angles

$$\left(0, 2\pi\sqrt{\frac{\log P}{n}}\right), \qquad \left(\pi - 2\pi\sqrt{\frac{\log P}{n}}, \pi + 2\pi\sqrt{\frac{\log P}{n}}\right),$$
$$\left(2\pi - 2\pi\sqrt{\frac{\log P}{n}}, 2\pi\right)$$

respectively. If H_1, H_2, H_3 denotes respectively, the number of roots of $f(z)$ in these angles we obtain

$$|H_1 - \sqrt{n \log P}| \leqq 16\sqrt{n \log P}, \qquad |H_2 - \sqrt{n \log P}| < 16\sqrt{n \log P}$$
$$|H_3 - \sqrt{n \log P}| < 16\sqrt{n \log P},$$

[9] Preuss. Akad. Wiss. Sitzungber. (1934), 99.

i.e. we obtain for the number H^* of the roots of $f(z)$ in the angles

$$|\operatorname{arc} z| < 2\pi\sqrt{\log P/n} \quad \text{and} \quad |\pi - \operatorname{arc} z| < 2\pi\sqrt{\log P/n}$$

(and *a fortiori* for the number H of the real roots) the estimation

$$H \leqq H^* < 51\sqrt{n \log P}$$

which is the theorem of Erhardt-Schmidt.

6. Now we turn to Jentzsch-Szegö's theorem. Let

$$f(z) = 1 + a_1 z + \cdots + a_n z^n + \cdots \tag{6.1}$$

be regular for $|z| < 1$ and let the unit-circle be the circle of convergence. Then for an arbitrary small positive ϵ we have an infinite sequence of indices,

$$n_1 < n_2 < \cdots$$

such that

$$|a_{n_\nu}| > (1 - \epsilon^2)^{n_\nu}, \qquad \nu = 1, 2, \cdots; \tag{6.2}$$

furthermore there is an $A_4 = A_4(\epsilon)$ such that for all $n > A_4(\epsilon)$ we have

$$|a_n| < (1 + \epsilon^2)^n. \tag{6.3}$$

We apply our Theorem I simply to the sections

$$s_{n_\nu}(z) = \sum_{j=0}^{n_\nu} a_j z^j. \tag{6.4}$$

In this case from (6.2) and (6.3) we have

$$P = \frac{1 + |a_1| + \cdots + |a_{n_\nu}|}{\sqrt{|a_{n_\nu}|}} < (1 - \epsilon^2)^{-(n_\nu/2)} \cdot \left(\sum_{n \leqq A_4(\epsilon)} |a_n| + \left(\frac{(1 + \epsilon^2)^{n_\nu+1} - 1}{\epsilon^2}\right)\right) < e^{9\epsilon^2 n_\nu},$$

if ϵ is sufficiently small and $n_\nu > A_5(\epsilon)$. For such n_ν's, denoting by $G_\nu(\alpha, \beta)$ the number of roots of $s_{n_\nu}(z)$ in the angle $\alpha \leqq \operatorname{arc} z \leqq \beta$, we obtain

$$\left| G_\nu(\alpha, \beta) - \frac{\beta - \alpha}{2\pi} n_\nu \right| < 48\epsilon n_\nu$$

for every α, β, with $0 \leqq \alpha < \beta \leqq 2\pi$.

In order to complete the proof of the theorem of Jentzsch-Szegö we have only to show that the number of roots of $s_{n_\nu}(z)$ outside the annulus

$$1 - \epsilon \leqq |z| \leqq 1 + \epsilon$$

is $< 7\,\epsilon n_\nu$ if $\nu > A_6(\epsilon)$. This is very simple. Let ϵ be so small that

$$\frac{1}{1 - \epsilon^2} < 1 + 2\epsilon^2 < e^{2\epsilon^2}. \tag{6.5}$$

Since the polynomials $s_{n_\nu}(z)$ converge uniformly in the circle $|z| \leqq 1 - \epsilon^2/2$, the number of roots in $|z| \leqq 1 - \epsilon^2$ (and thus *a fortiori* the number of roots in $|z| \leqq 1 - \epsilon$) for $\nu > A_7(\epsilon)$ does not exceed $2A_8(\epsilon)$, where $A_8(\epsilon)$ denotes the number of roots of $f(z)$ in the circle $|z| \leqq 1 - \frac{1}{2}\epsilon^2$. Further, since $s_{n_\nu}(0) = 1$, denoting by $z_j(j = 1, \cdots, n_\nu)$ the roots of $s_{n_\nu}(z)$ we have

$$\prod_{|z_j| \leqq 1-\epsilon^2} |z_j| > A_9(\epsilon). \tag{6.6}$$

Now from (6.2) for these n_ν's

$$(1 - \epsilon^2)^{n_\nu} \leqq |a_{n_\nu}| = \prod_{j=1}^{n} |z_j|^{-1} = \prod_{|z_j| \leqq 1-\epsilon^2}{}_j \; \prod_{1-\epsilon^2 \leqq |z_j| \leqq 1+\epsilon}{}_j \; \prod_{|z_j| > 1+\epsilon} . \tag{6.7}$$

Denoting by J_ν the number of roots of $s_{n_\nu}(z)$ in $|z| \geqq 1 + \epsilon$ we obtain from (6.6) and (6.7)

$$(1 - \epsilon^2)^{n_\nu} \leqq A_9(\epsilon)^{-1} \left(\frac{1}{1 - \epsilon^2}\right)^{n_\nu - J_\nu} \left(\frac{1}{1 + \epsilon}\right)^{J_\nu}, \tag{6.8}$$

$$\left(\frac{1}{1 - \epsilon}\right)^{J_\nu} \leqq A_9(\epsilon)^{-1} \left(\frac{1}{1 - \epsilon^2}\right)^{2n_\nu}.$$

Since for $\nu > A_{10}(\epsilon)$ we have

$$2A_8(\epsilon) < \epsilon n_\nu, \qquad \frac{1}{A_9(\epsilon)} < \left(\frac{1}{1 - \epsilon^2}\right)^{n_\nu}, \tag{6.9}$$

choosing

$$A_{11}(\epsilon) = \max\,(A_{10}(\epsilon), A_6(\epsilon), A_7(\epsilon))$$

we have from (6.8), (6.9) and (6.5) for $\nu > A_{11}(\epsilon)$

$$e^{\epsilon J_\nu} < \left(\frac{1}{1 - \epsilon}\right)^{J_\nu} < \left(\frac{1}{1 - \epsilon^2}\right)^{3n_\nu} < e^{6\epsilon^2 n_\nu}.$$

From this and (6.9) we have for the total number of roots of $s_{n_\nu}(z)$ outside of the annulus $1 - \epsilon \leqq |z| \leqq 1 + \epsilon$ the upper estimation

$$< 6\epsilon n_\nu + 2A_8(\epsilon) < 7\epsilon n_\nu . \qquad \text{Q.e.d.}$$

7. If we know something about the coefficients of the power series (6.1), then in a similar way we can obtain more exact information about the distribution of the roots of the sections. We formulate only

THEOREM II. *If for the coefficients of the power-series* (6.1) *we have*

$$\nu^{-\lambda} \leqq |a_\nu| \leqq \nu^{\lambda}, \qquad \nu = 1, 2, \cdots \tag{7.1}$$

then there is a $A_{12} = A_{12}(\lambda)$ *such that for the roots* $z_1, z_2, \cdots z_n$ *of the section* $s_n(z)$ *we have for any* $0 \leqq \alpha < \beta \leqq 2\pi$

$$\left| \sum_{\substack{\alpha \leqq \arc z \leqq \beta \\ 1-(1/\sqrt{n}) \leqq |z_j| \leqq 1+(1/\sqrt{n})}}{}_j 1 - \frac{\beta - \alpha}{2\pi} n \right| < A_{12}(\lambda) \sqrt{n \log n}.$$

The proof goes along the same lines as in §6 so we can omit the details.

8. We shall prove our Theorem I combining the method of our joint paper[10] (suitably modified) with an artifice of Schur[4] in §§10–14. In our paper[10] we used the mentioned method to prove the theorem that if

$$\omega_n(x) = x^n + \cdots + a_n \tag{8.1}$$

has all its roots in $-1 \leqq x \leqq +1$ and satisfies here the inequality[11]

$$| \omega_n(x) | \leqq B(n)/2^n, \tag{8.2}$$

then, writing the roots in the form

$$\cos \vartheta_\nu, \qquad \nu = 1, 2, \cdots, n, \qquad 0 \leqq \vartheta_1 \leqq \vartheta_2 \leqq \cdots \leqq \vartheta_n \leqq \pi,$$

we have for all α, β with $0 \leqq \alpha < \beta \leqq \pi$

$$\left| \sum_{\alpha \leqq \vartheta_\nu \leqq \beta}{}_{\nu} 1 - \frac{\beta - \alpha}{\pi} n \right| < \frac{8}{\log 3} \sqrt{n \log B(n)}. \tag{8.3}$$

We applied this estimation[10] to obtain rather exact results concerning the distribution of the roots of some polynomials which are defined by minimizing properties. More exactly, given a $k > 0$ and a function $p(x)$ non-negative and L-integrable in $-1 \leqq x \leqq +1$ we have, as Dunham Jackson[12] proved, a sequence of polynomials

$$\varphi_0(x, k), \varphi_1(x, k), \cdots, \varphi_n(x, k), \cdots \tag{8.4}$$
$$\varphi_n(x, k) = x^n + \cdots + b_n$$

such that $\varphi_n(x, k)$, minimizes the integral

$$\int_{-1}^{1} | \pi_n(x) |^k p(x) \, dx \tag{8.5}$$

among the polynomials $\pi_n(x) = x^n + \cdots$. If $k = 2$ the polynomials are identical with the orthogonal polynomials belonging to the weight function $p(x)$; the properties of these polynomials have been extensively studied and one can find the whole theory e.g. in Szegö's[13] book Orthogonal Polynomials. However, for general $k > 0$ the corresponding theory does not exist; in our paper[10] we could settle one of the four main problems of the theory, the uniform distribution of the roots on the segment $-1 \leqq x \leqq +1$ in the rather general case when the weight function $p(x)$ satisfies the condition

$$p(x) \geqq A_{13} > 0. \tag{8.6}$$

[10] Ann. of Math. (2) 41 (1940), 162–173. We shall use this opportunity to correct a number of misprints in the paper. In the first chain of inequalities (between (1) and (2)) and in inequality (2) the quantity π is to be replaced by 2π and in all occurrences on this page the sign $\leqq$ is replaced by $<$. In the footnote of page 163 the letter ρ is replaced by l. On page 167 in the fourth line from the bottom $[\xi_{\nu+2}^{(n)}, \xi_\nu^{(n)}]$ is replaced by $[\xi_{\nu+1}^{(n)}, \xi_\nu^{(n)}]$.

[11] According to the classical theorem of Chebyshev we have always $B(n) \geqq 2$.

[12] Trans. Amer. Math. Soc. 22 (1921), 117–128; 320–326. As he proved in another paper, ibid. 25 (1923), 333-338, in the case $0 < k < 1$ the unicity can fail.

[13] Amer. Math. Soc. Colloquium Publications, Vol. 23, 1939.

We proved namely that in this case, denoting the roots of $\varphi_n(x, k)$ by $\cos \vartheta_\nu$ with

$$0 \leqq \vartheta_1 \leqq \vartheta_2 \leqq \cdots \leqq \vartheta_n \leqq \pi,$$

we have for all $0 \leqq \alpha < \beta \leqq \pi$

$$\left| \sum_{\alpha \leqq \vartheta_\nu \leqq \beta}^{\nu} 1 - \frac{\beta - \alpha}{\pi} n \right| \leqq A_{14}(k, p) \sqrt{n \log n}. \tag{8.7}$$

Now in an Appendix we shall show by simple reasoning how from (8.7) we can obtain a partial solution of another main problem, the problem of "outer"-asymptotic representation of the polynomials $\varphi_n(z, k)$. By this we mean an asymptotic representation valid on every point of the complex z-plane cut along the segment $-1 \leqq z \leqq +1$. In the case of $k = 2$ Szegö[14] proved, even in the more general case when instead of (8.4) we suppose only that $\log p(x)$ is L-integrable, the asymptotical representation

$$\varphi_n(z, 2) = \left(\frac{z + \sqrt{z^2 - 1}}{2}\right)^{n + J(z) + o(1)}$$

where the meaning of $\sqrt{z^2 - 1}$ is obvious, the function $J(z)$ is determned by the weight function $p(x)$, the o-sign refers to $n \to \infty$ and holds uniformly in every domain not having a common point with the interval $[-1, +1]$. In the case of $k > 0$ and $p(x)$ satisfying (8.6) we shall prove the asymptotical representation

$$\varphi_n(z, k) = \left(\frac{z + \sqrt{z^2 - 1}}{2}\right)^{n + 0(\sqrt{n \log n})} \tag{8.8}$$

where the 0-sign holds again uniformly in every domain not having a common point with the interval $[-1, +1,]$. Using a much more difficult argument P. Erdös succeeded in improving the error-term in (8.7) to $0(\log^2 n)$-which is certainly not far from the best possible and also that in the exponent in (8.8) to $0(\log^2 n)$. The asymptotical representation (8.8) has the following consequence. If $F(z)$ is regular along $[-1, +1]$, then expanding $F(z)$ in a series

$$F(z) \sim \sum_{\nu=0}^{\infty} c_\nu \varphi_\nu(z, k)$$

the domain of convergence will be an ellipse with the points ± 1 as foci. This was well-known in the case $k = 2$.

9. Before turning to the proof of Theorem I, we consider the polynomials

$$h(x) = \sum_{p \leqq n} z^p$$

where p runs over the prime numbers not exceeding n. These polynomials play an essential part in the additive theory of primes and recent investigations concerning Riemann's famous hypothesis[15] carried out by P. Turán attach some

[14] See e.g. his book[13] p. 290, theorem 12.1.2.

[15] Unpublished.

importance to any information one can obtain about their roots. Our Theorem I gives obviously the estimation

$$\left| V(\alpha, \beta, h) - \frac{\beta - \alpha}{2\pi} n \right| < 16\sqrt{n \log n}$$

where $V(\alpha, \beta, h)$ denotes the number of roots of $h(z)$ in the sector

$$\alpha \leqq \operatorname{arc} z \leqq \beta.$$

10. Now we turn to the proof of our Theorem I. From (3.1) and (3.2) we have

$$f(z) = a_n \prod_{\nu=1}^{n} (z - z_\nu). \tag{10.1}$$

We introduce the polynomial $g(z)$ by

$$g(z) = \prod_{\nu=1}^{n} (z - e^{i\varphi_\nu}). \tag{10.2}$$

Let $z = e^{i\varphi}$ be fixed on the unit-circle and let $\xi = \rho e^{i\vartheta}$ run on the ray arc $z = \vartheta$, ϑ fixed. Then

$$| z - \xi |^2 = 1 + \rho^2 - 2\rho \cos (\varphi - \vartheta)$$

$$\frac{| z - \xi |^2}{|\xi|} = \rho + 1/\rho - 2 \cos (\varphi - \vartheta).$$

But if ρ varies from 0 to $+\infty$, the right side has its minimum at $\rho = 1$; hence

$$\frac{| z - \xi |^2}{|\xi|} \geqq 2 - 2 \cos (\varphi - \vartheta) = | z - e^{i\vartheta} |^2.$$

This is Schur's previously mentioned remark. Applying this with

$$\zeta = z_\nu, \qquad \nu = 1, 2, \cdots, n$$

and multiplying we obtain on the unit-circle

$$\frac{| \prod_{\nu=1}^{n} (z - z_\nu) |^2}{| \prod_{\nu=1}^{n} z_\nu |} \geqq | \prod_{\nu=1}^{n} (z - e^{i\varphi_\nu}) |^2.$$

i.e.

$$| g(z) |^2 \leqq \frac{|f(z)|^2}{| a_n |^2 | \prod_{\nu=1}^{n} z_\nu |} = \frac{|f(z)|^2}{| a_0 a_n |} \leqq \left(\frac{| a_0 | + \cdots + | a_n |}{\sqrt{| a_0 a_n |}} \right)^2 = P^2$$

or in the whole unit-circle

$$| g(z) | \leqq P. \tag{10.3}$$

Now we fix α, β satisfying

$$0 \leqq \alpha < \beta \leqq 2\pi$$

and denote by $H(\alpha, \beta, g)$ the number of φ_ν satisfying

$$\alpha \leqq \varphi_\nu \leqq \beta. \tag{10.4}$$

If we can prove

$$\begin{aligned} -16 \sqrt{n \log \max_{|z|=1} |g(z)|} \leqq H(\alpha, \beta, g) \\ - \frac{\beta - \alpha}{2\pi} n \leqq 16 \sqrt{n \log \max_{|z|=1} g(z) |}, \end{aligned} \tag{10.5}$$

then owing to (10.3) our Theorem I will be proved. Let us suppose we have proved the sharper upper estimation

$$H(\gamma, \delta, g) \leqq \frac{\delta - \gamma}{2\pi} n + 8 \sqrt{n \log \max_{|z|=1} |g(z)|} \tag{10.6}$$

for every $0 \leqq \gamma < \delta \leqq 2\pi$. Then applying it once with

$$\gamma = 0, \qquad \delta = \alpha$$

then with

$$\gamma = \beta, \qquad \delta = 2\pi$$

we obtain

$$\begin{aligned} H(\alpha, \beta, g) = n - H(0, \alpha, g) - H(\beta, 2\pi, g) \geqq \frac{\beta - \alpha}{2\pi} n - 8 \sqrt{n \log \max_{|z|=1} |g(z)|} \\ - 8 \sqrt{n \log \max_{|z|=1} |g(z)|} = \frac{\beta - \alpha}{2\pi} n - 16 \sqrt{n \log \max_{|z|=1} |g(z)|} \end{aligned}$$

i.e. also the lower estimation of (10.5) could be proved. Hence in order to prove our Theorem I it is sufficient to prove inequality (10.6). Obviously we may suppose without loss of generality $\gamma = 0$.

11. Fixing the number of roots of $g(z)$ on the arc

$$0 \leqq \varphi \leqq \delta$$

of the unit-circle, the inequality (10.6) gives a lower estimation of $\max | g(z) |$. Hence we shall consider the following extremal problem of Tshebisheff-type:

Using the abbreviation

$$\left[\frac{\delta}{2\pi} n\right] = K, \tag{11.1}$$

what is the minimal value M of the absolute maximum on $| z | = 1$ of those polynomials $g(z)$ of the form

$$g(z) = \prod_{\nu=1}^{n} (z - e^{i\vartheta_\nu}), \qquad | c | = 1, \qquad 0 \leqq \vartheta_1 \leqq \vartheta_2 \leqq \cdots \leqq \vartheta_n < \vartheta\pi, \tag{11.2}$$

which have exactly $(K + 2l + 1)$ roots on the arc $0 \leqq \varphi \leqq \delta$?

12. As is well known, such a minimum exists. We shall call all polynomials of our class given in (11.2) whose absolute maximum is M, extremal-polynomials. Let $g_0(z)$ be one such extremal-polynomial. First we assert the following:

LEMMA. *$g_0(z)$ assumes on every arc $[\vartheta_\nu, \vartheta_{\nu+1}]$ of the unit-circle which lies inside the arc $0 \leqq \varphi \leqq \delta$ a value which is $= M$ in absolute value.*

PROOF. Suppose the lemma is not true; we may suppose that on the arc $\vartheta_1 \leqq \varphi \leqq \vartheta_2$ of the unit-circle we have

$$|g_0(z)| \leqq M - \epsilon_0, \qquad \epsilon_0 > 0 \tag{12.2}$$

(Of course the same holds if we replace ϵ_0 by an arbitrary positive $\epsilon_1 \leqq \epsilon_0$.) Owing to the continuity of $g_0(z)$ we can find a positive ϵ_2 such that on the larger arc

$$\vartheta_1 - \epsilon_2 \leqq \varphi \leqq \vartheta_2 + \epsilon_2 \tag{12.3}$$

we have

$$|g_0(z)| \leqq M - \frac{\epsilon_0}{2}. \tag{12.4}$$

Owing to the continuous dependence of the coefficients on the zeros we can choose ϵ_3 so small that for every positive $\eta < \epsilon_3$ the polynomial

$$g_1(z) = \frac{g_0(z)}{(z - e^{i\vartheta_1})(z - e^{i\vartheta_2})}(z - e^{i(\vartheta_1 - \eta)})(z - e^{i(\vartheta_2 + \eta)}) \tag{12.5}$$

should be absolutely $\leqq M - \epsilon_0/3$ on the arc (12.3). Summarizing the restrictions on η we have

$$0 < \eta < \min(\vartheta_1, \delta - \vartheta_2, \epsilon_2, \epsilon_3)$$

and then $g_1(z)$ belongs certainly to our class (11.2).

We estimate $g_1(z)$ on the complementary arc $\bar{I}$ of (12.3) after we know that on the arc (12.3) itself

$$|g_1(z)| \leqq M - \frac{\epsilon_0}{3}. \tag{12.6}$$

Since for small η on the $\bar{I}$

$$\frac{z - e^{i(\vartheta_1 - \eta)}}{z - e^{i\vartheta_1}} = 1 + \eta i \frac{e^{i\vartheta_1}}{z - e^{i\vartheta_1}} + 0(\eta^2)$$

$$\frac{z - e^{i(\vartheta_2 + \eta)}}{z - e^{i\vartheta_2}} = 1 - \eta i \frac{e^{i\vartheta_2}}{z - e^{i\vartheta_2}} + 0(\eta^2)$$

we have here

$$g_1(z) = g_0(z)\left\{1 + i\eta z \frac{e^{i\vartheta_1} - e^{i\vartheta_2}}{(z - e^{i\vartheta_1})(z - e^{i\vartheta_2})} + 0(\eta^2)\right\}. \tag{12.7}$$

But putting $z = e^{i\varphi}$ we have

$$1 + i\eta z \frac{e^{i\vartheta_1} - e^{i\vartheta_2}}{(z - e^{i\vartheta_1})(z - e^{i\vartheta_2})} = 1 - \eta \frac{\sin \frac{\vartheta_2 - \vartheta_1}{2}}{2 \sin \frac{\vartheta - \vartheta_1}{2} \sin \frac{\vartheta - \vartheta_2}{2}}.$$

Since the coefficient of η is on $\bar{I}$ greater than a *positive* quantity independent of η, choosing η sufficiently small, the bracketed expression in (12.7) is less than 1 in absolute value on $\bar{I}$ and hence on this arc

$$| g_1(z) | < M.$$

But this and (12.6) constitute a contradiction in view of the fact that M was supposed to be the *minimum* of the absolute maxima of the polynomials of our class, and our lemma is proved.

13. Next we apply the following theorem[16] due to P. Turán.

If an arbitrary polynomial of degree n assumes its absolute maximum on $| z | = 1$ for $z = e^{i\varphi_0}$ then the arc

$$| \varphi_0 - \text{arc } z | < \frac{\pi}{n}$$

of the unit-circle is free of roots of this polynomial.

The application of this theorem—taking into account our lemma—shows that the distance of any two consecutive roots *inside* the arc $0 \leqq \text{arc } z \leqq \delta$ of the unit-circle is $\geqq 2\pi/n$; hence the inside of this arc can contain only $1 + [(\delta/2\pi)n] = 1 + K$ roots at most. Owing to the definition of our class at least $2l$ roots must be at the end-points of the arc $0 \leqq \varphi \leqq \delta$ and thus *the extremal-polynomials have at least one root with the multiplicity* l. Then without loss of generality

$$g_0(z) = c(1 + z)^l g_2(z) \tag{13.1}$$

$$| c | = 1, \qquad g_2(z) = z^{n-l} + d_1 z^{n-l-1} + \cdots + d_{n-l}.$$

Further,

$$\begin{aligned} M^2 &\geqq \frac{1}{2\pi} \int_{|z|=1} | g_0(z) |^2 \, | dz | = \frac{1}{2\pi} \int_{|z|=1} | 1 + z |^{2l} \cdot | g_2(z) |^2 \, | dz | \\ &\geqq \min_G \frac{1}{2\pi} \int_{|z|=1} | 1 + z |^{2l} \cdot | g_2(z) |^2 \, | dz |, \end{aligned} \tag{13.2}$$

where class G means the polynomials of degree $(n - l)$ with the leading coefficient 1.

14. We now determine the last minimum exactly. According to a general theorem of Szegö[17] the integral (13.2) attains its minimum for one polynomial

[16] Szeged Acta 11 (1946), 106–113.

[17] E.g. his Orthogonal Polynomials, p. 282, theorem 11.1.2.

$g_3(z)$ of the class G which is characterised by the conditions

$$(14.1)\quad \int_{|z|=1} g_3(z)\,|\,1 + z\,|^{2l}(\bar{z})^{\nu}\,|\,dz\,| = 0, \qquad \nu = 0, 1, \cdots (n - l - 1),$$

or equivalently

$$(14.2)\qquad \int_{|z|=1} g_3(z)\,\frac{(1 + z)^{2l}}{z^{l+\nu+1}}\,dz = 0, \qquad \nu = 0, 1, \cdots (n - l - 1).$$

Now we assert that

$$(14.3)\quad g_3(z) = \frac{l\binom{n + l}{l}}{(1 + z)^{2l}} \int_{-1}^{z} (z - t)^{l-1}(1 + t)^{l}\,t^{n-l}\,dt \equiv \frac{l\binom{n + l}{l}}{(1 + z)^{2l}}\,\psi(z).$$

For $0 \leqq \nu \leqq l - 1$ we have

$$\psi^{(\nu)}(z) = \frac{(l - 1)!}{(l - \nu - 1)!} \int_{-1}^{z} (z - t)^{l-\nu-1}(1 + t)^{l}\,t^{n-l}\,dt$$

i.e.

$$(14.4)\qquad \psi^{(\nu)}(-1) = 0, \qquad \nu = 0, 1, \cdots, (l - 1).$$

Further we have

$$\psi^{(l)}(z) = (l - 1)!(1 + z)^{l}z^{n-l},$$

i.e.

$$(14.5)\qquad \psi^{(\nu)}(-1) = 0, \qquad \nu = l, (l + 1), \cdots, (2l - 1).$$

From (14.4) and (14.5) we see that $\psi(z)$ has a zero of order $2l$ at least for $z = -1$; hence $g_3(z)$ defined by (14.3) is really a polynomial of degree $(n - l)$. Further

$$\begin{aligned}\text{coeffs. } z^{n-l} \text{ in } g_3(z) &= l\binom{n + l}{l} \lim_{z\to\infty} \frac{1}{z^{n+l}} \int_{-1}^{z} (z - t)^{l-1}(1 + t)^{l}\,t^{n-l}\,dt\\ &= l\binom{n + l}{l} \lim_{z\to\infty} \frac{1}{z^{n+l}} \int_{-1}^{z} (z - t)^{l-1}\,t^{n}\,dt\\ &= l\binom{n + l}{l} \lim_{z\to\infty} \int_{-1/z}^{1} (1 - y)^{l-1}\,y^{n}\,dy = 1;\end{aligned}$$

hence $g_3(z)$ really belongs to our class G. We have to verify (14.2). Using (14.3) we have only to verify

$$\int_{|z|=1} \frac{\psi(z)}{z^{l+\nu+1}}\,dz = 0, \qquad \nu = 0, 1, \cdots, (n - l - 1)$$

i.e.,

$$\text{Coeffs } z^{l} = \text{coeffs } z^{l+1} = \cdots = \text{coeffs. } z^{n-1} = 0$$

in $\psi(z)$. Writing $\psi(z)$ in the form

$$(14.6)\quad \psi(z) = \int_{-1}^{0} (z - t)^{l-1}(1 + t)^l t^{n-l}\, dt + \int_0^z (z - t)^{l-1}(1 + t)^l t^{n-l}\, dt,$$

we see that the first integral represents a polynomial of degree $(l - 1)$ and the second can be written after the substitution $t = zy$ in the form

$$(14.7)\qquad z^n \int_0^1 (1 - y)^{l-1}(1 + zy)^l y^{n-l}\, dy$$

which is a polynomial in which the *smallest* exponent is n. Hence really $g_3(z)$ is the polynomial minimizing the integral (13.2) in the class G.

We can calculate easily the value of this minimum.

$$\frac{1}{2\pi}\int_{|z|=1} |g_3(z)|^2 \, |1 + z|^{2l} \, |dz| = \frac{1}{2\pi}\int_{|z|=1} g_3(z)(\bar{z})^{n-l} \, |1 + z|^{2l} \, |dz|$$

$$= \frac{1}{2\pi i}\int_{|z|=1} \frac{g_3(z)(1 + z)^{2l}}{z^{n+1}}\, dz = \frac{1}{2\pi i}\, l \binom{n + l}{l} \int_{|z|=1} \frac{\psi(z)}{z^{n+1}}\, dz.$$

Taking into account (14.6) and (14.7), the value of this minimum is

$$l \binom{n + l}{l} \int_0^1 (1 - y)^{l-1} y^{n-l}\, dy = \frac{\binom{n + l}{l}}{\binom{n}{l}},$$

i.e. from (13.2)

$$M^2 \geqq \frac{\binom{n + l}{l}}{\binom{n}{l}}$$

and for all polynomials of our class (11.2)

$$\max_{|z|=1} |g(z)|^2 \geqq \frac{\binom{n + l}{l}}{\binom{n}{l}} = \prod_{\nu=n-l+1}^{n} \left(1 + \frac{l}{\nu}\right),$$

$$2 \log \max_{|z|=1} |g(z)| \geqq \sum_{\nu=n-l+1}^{n} \frac{\frac{l}{\nu}}{1 + \frac{l}{\nu}} = l \sum_{\nu=n+1}^{n+l} \frac{1}{\nu}$$

$$> l \log\left(1 + \frac{l}{n + 1}\right) > l\, \frac{\frac{l}{n + 1}}{1 + \frac{l}{n + 1}} \geqq \frac{l^2}{2(n + 1)},$$

$$l < 2\sqrt{(n + 1) \log \max_{|z|=1} |g(z)|} \leqq 2\sqrt{(n + 1) \log P}.$$

Hence, if P is fixed, the number of roots in the angle $0 \leqq \varphi \leqq \delta$ is

$$< \left[\frac{\delta}{2\pi} n\right] + 4\sqrt{(n+1)\log P} + 1. \tag{14.8}$$

Since

$$\frac{|a_0| + \cdots + |a_n|}{\sqrt{|a_0 a_n|}} \geqq \frac{|a_0| + |a_n|}{\sqrt{|a_0 a_n|}} \geqq 2$$

we have

$$4\sqrt{(n+1)\log P} + 1 < 8\sqrt{n \log P}$$

which together with (14.8) proves the required inequality (10.6).

Appendix

As we mentioned we shall establish for the polynomials (8.4) minimizing the integral (8.5) under the restriction (8.6) the asymptotical representation (8.8).

Designating the roots of the minimizing polynomials (8.4) (which all lie according to the theorem of Fejér[18] in the interval $-1 \leqq x \leqq +1$) by $\cos \vartheta_\nu$ where

$$0 \leqq \vartheta_1 \leqq \vartheta_2 \leqq \cdots \leqq \vartheta_n \leqq \pi,$$

we first show that for $\nu = 1, 2, \cdots, n$ we have

$$\vartheta_\nu \leqq \frac{2\nu - 1}{2n}\pi + 2\pi A_{14}(k, p)\sqrt{\frac{\log n}{n}} \tag{15.1}$$

($A_{14}(k, p)$ defined in (8.7)). For if this were not true for $\nu = \nu_0$ we would have

$$\vartheta_{\nu_0} > \frac{2\nu_0 - 1}{2n}\pi + 2\pi A_{14}(k, p)\sqrt{\frac{\log n}{n}},$$

i.e. the interval

$$0 \leqq \vartheta \leqq \frac{2\nu_0 - 1}{2n}\pi + 2\pi A_{14}(k, p)\sqrt{\frac{\log n}{n}}$$

would contain less than ν_0 of the ϑ's. But applying the inequality (8.7) with $\alpha = 0, \beta = (2\nu_0 - 1)/2n\,\pi + 2\pi A_{14}(k, p)\sqrt{\log n/n}$ gives

$$\left|\sum_{0 \leqq \vartheta_j \leqq (2\nu_0 - 1/2n)\pi + 2\pi A_{14}(k,p)\sqrt{\log n/n}} 1 - \frac{2\nu_0 - 1}{2} - 2A_{14}(k, p)\sqrt{n \log n}\right| < A_{14}(k, p)\sqrt{n \log n},$$

i.e.

$$\left|\sum_{0 \leqq \vartheta_j \leqq (2\nu_0 - 1/2n)\pi + 2\pi A_{14}(k,p)\sqrt{\log n/n}} 1\right| > \nu_0 - \tfrac{1}{2} + A_{14}(k, p)\sqrt{n \log n} > \nu_0$$

[18] Math. Ann. 85 (1922), 41–48.

which is a contradiction. Similarly we can show that

$$\vartheta_\nu \geqq \frac{2\nu - 1}{2n}\pi - 2\pi A_{14}(k, p)\sqrt{\frac{\log n}{n}} \tag{15.2}$$

$$\nu = 1, 2, \cdots, n.$$

From (15.1) and (15.2) if

$$-D \equiv -(\pi A_{14}(k, p)^2 + 2\pi A_{14}(k, p)) \leqq \lambda_\nu \leqq D$$

$$\nu = 1, 2, \cdots, n$$

we have

$$\varphi_n(z, k) = \prod_{\nu=1}^{n}\left(z - \cos\frac{2\nu - 1}{2n}\pi + \lambda_\nu\sqrt{\frac{\log n}{n}}\right)$$

$$= \prod_{\nu=1}^{n}\left(z - \cos\frac{2\nu - 1}{2n}\pi\right)\exp\left(\vartheta D\sqrt{\frac{\log n}{n}}\sum_{\nu=1}^{n}\frac{1}{\left|z - \cos\frac{2\nu - 1}{2n}\pi\right|}\right)$$

with $|\vartheta| \leqq 1$. Hence if

$$\min_{-1\leqq x\leqq +1}|z - x| = E > 0$$

then

$$\varphi_n(z, k) = \prod_{\nu=1}^{n}\left(z - \cos\frac{2\nu - 1}{2n}\pi\right)e^{\vartheta'(D/E)\sqrt{n\log n}} \tag{15.3}$$

with $|\vartheta'| \leqq 1$. But

$$\prod_{\nu=1}^{n}\left(z - \cos\frac{2\nu - 1}{2n}\pi\right) = T_n(z) = \left(\frac{z + \sqrt{z^2 - 1}}{2}\right)^n + \left(\frac{z - \sqrt{z^2 - 1}}{2}\right)^n$$

$$= \left(\frac{z + \sqrt{z^2 - 1}}{2}\right)^n[1 + (z - \sqrt{z^2 - 1})^{2n}] = \left(\frac{z + \sqrt{z^2 - 1}}{2}\right)^n(1 + \vartheta''^n)$$

with $|\vartheta''| \leqq 1$. This with (15.3) proves our assertion.

THE INSTITUTE FOR ADVANCED STUDY