COMPOSITIO MATHEMATICA

I. J. SCHOENBERG
GABOR SZEGÖ

An extremum problem for polynomials

Compositio Mathematica, tome 14 (1959-1960), p. 260-268
<http://www.numdam.org/item?id=CM_1959-1960__14__260_0>

An Extremum Problem for Polynomials [1]

by

I. J. Schoenberg and Gabor Szegö

1. Introduction

Let $[a, b]$ be a finite interval and let the measure $d\chi(x)$ define a positive mass-distribution which does not reduce to a finite number of point-masses. The natural integer n being given, we denote by Π_n the class of polynomials $f(x)$ of degree not exceeding n and subject to the following two conditions:

$$f(x) \geqq 0 \text{ in } [a, b], \tag{1}$$

$$\int_a^b f(x) d\chi(x) = 1. \tag{2}$$

Now z being an arbitrary but fixed real number, we wish to determine the range R_z of variability of $f(z)$ as the polynomial $f(x)$ roams through the class Π_n defined above. Since Π_n is a convex class, it is clear that R_z is convex, hence an interval which is clearly also bounded and closed. It may therefore be described by

$$R_z = [\min_{f \in \Pi_n} f(z), \max_{f \in \Pi_n} f(z)]. \tag{3}$$

On page 181 of his book [3] Szegö determines the quantity

$$\max_{f \in \Pi_n} | f(z)|.$$

We reproduce here his brief analysis which also readily describes the interval (3). Let

$$\{p_\nu(x)\}, \{q_\nu(x)\}, \{r_\nu(x)\}, \{s_\nu(x)\}, \tag{4}$$

be the orthonormal sets of polynomials associated with the mass-distributions

$$d\chi(x), (x-a)(b-x)d\chi(x), (x-a)d\chi(x), (b-x)d\chi(x), (a \leqq x \leqq b). \tag{5}$$

respectively. The leading coefficients of the polynomials (4) are assumed to be positive. An important theorem of Lukács [3, p. 4] allows to describe the most general $f(x) \in \Pi_n$ as

---

[1]) This paper was prepared partly under the sponsorship of the United States Air Force, Office of Scientific Research, ARDC, and partly under the sponsorship of the National Science Foundation, NSF-G 11296.

$$f(x) = (\sum_0^k u_\nu p_\nu(x))^2 + (x-a)(b-x)(\sum_0^{k-1} v_\nu q_\nu(x))^2 \quad \text{if } n = 2k, \tag{6}$$

or

$$f(x) = (x-a)(\sum_0^k u_\nu r_\nu(x))^2 + (b-x)(\sum_0^k v_\nu s_\nu(x))^2 \quad \text{if } n = 2k+1, \tag{7}$$

where the u_ν, v_ν are real parameters satisfying the single relation

$$\sum_0^k u_\nu^2 + \sum_0^k v_\nu^2 = 1, \quad (v_k = 0 \text{ if } n = 2k). \tag{8}$$

The relations (6), (7) express the condition (1), while (8) is equivalent to the normalization (2) as seen from the orthogonality properties of the sets (4).

Let us use the abbreviations

$$\begin{aligned} \Phi_1(x) &= \sum_0^k (p_\nu(x))^2, & \Phi_2(x) &= (x-a)(b-x)\sum_0^{k-1}(q_\nu(x))^2 \\ \Psi_1(x) &= (x-a)\sum_0^k(r_\nu(x))^2, & \Psi_2(x) &= (b-x)\sum_0^k(s_\nu(x))^2. \end{aligned} \tag{9}$$

Applying Cauchy's inequality to one or to both terms on the right-hand sides of (6) and (7), as required by the location of $x = z$, we readily obtain the range R_z as stated in the following

THEOREM 1. *If $n = 2k$ then*

$$\min_{f \in \Pi_n} f(z) = \begin{cases} 0 & \text{if } a \leqq z \leqq b, \ (k \geqq 1), \\ \Phi_2(z) & \text{if } z < a \text{ or } z > b, \end{cases} \tag{10}$$

$$\max_{f \in \Pi_n} f(z) = \begin{cases} \max\{\Phi_1(z), \ \Phi_2(z)\} & \text{if } a \leqq z \leqq b, \\ \Phi_1(z) & \text{if } z < a \text{ or } z > b. \end{cases} \tag{11}$$

If $n = 2k+1$ then

$$\min_{f \in \Pi_n} f(z) = \begin{cases} 0 & \text{if } a \leqq z \leqq b, \ (k \geqq 1), \\ \Psi_1(z) & \text{if } z < a, \\ \Psi_2(z) & \text{if } z > b, \end{cases} \tag{12}$$

$$\max_{f \in \Pi_n} f(z) = \begin{cases} \max\{\Psi_1(z), \ \Psi_2(z)\} & \text{if } a \leqq z \leqq b, \\ \Psi_1(z) & \text{if } z > b, \\ \Psi_2(z) & \text{if } z < a. \end{cases} \tag{13}$$

2. Two identities and their applications

We see that $\min f(z)$ is well described by the formulae (10) and (12). The purpose of this note is to improve on the description of $\max f(z)$ as given by the formulae (11), (13), for the case when z is in $[a, b]$. This will be done by exhibiting all the crossing points of the graphs of $\Phi_1(z)$, $\Phi_2(z)$ and also those of the graphs of $\Psi_1(z)$ and $\Psi_2(z)$.

These crossings are described by

LEMMA 1. *The following identities hold*:

$$(14)\qquad \sum_0^k (p_\nu(x))^2-(x-a)(b-x)\sum_0^{k-1}(q_\nu(x))^2 = A_k r_k(x)s_k(x),$$

$$(15)\qquad (x-a)\sum_0^k (r_\nu(x))^2-(b-x)\sum_0^k (s_\nu(x))^2 = B_k p_{k+1}(x)q_k(x),$$

where A_k, B_k *are positive constants.*

Let us establish the identity (14). As a matter of fact the more general identity in the two independent variables x and ξ holds:

$$(16)\quad \sum_0^k p_\nu(\xi)p_\nu(x)-(x-a)(b-\xi)\sum_0^{k-1}q_\nu(\xi)q_\nu(x) = A_k r_k(\xi)s_k(x).$$

It might be described as arising from (14) by "polarization." To show (16) we use the reproducing property of the kernel polynomial

$$K(\xi,\ x) = \sum_0^k p_\nu(\xi)p_\nu(x)$$

which is as follows [3, pp. 38—39]: If $\pi_k(x)$ is an arbitrary polynomial of degree k then

$$\int_a^b K(\xi,\ x)\pi_k(x)d\chi(x) = \pi_k(\xi).$$

Also the other kernel polynomials

$$\sum_0^{k-1} q_\nu(\xi)q_\nu(x),\ \sum_0^k r_\nu(\xi)r_\nu(x),\ \sum_0^k s_\nu(\xi)s_\nu(x),$$

have similar properties with respect to their corresponding mass-distributions. Denoting by $\Phi(\xi, x)$ the left-hand side of (16) we find

$$\int_a^b \Phi(\xi,\ x)(b-x)\pi_{k-1}(x)d\chi(x) = (b-\xi)\pi_{k-1}(\xi)-(b-\xi)\pi_{k-1}(\xi) = 0.$$

But then

$$(17)\qquad \Phi(\xi,\ x) = C_1(\xi)s_k(x).$$

Similarly

$$\int_a^b \Phi(\xi,\ x)(\xi-a)\pi_{k-1}(\xi)d\chi(\xi) = (x-a)\pi_{k-1}(x)-(x-a)\pi_{k-1}(x) = 0$$

and therefore

$$(18) \qquad \Phi(\xi, x) = C_2(x) r_k(\xi).$$

Clearly (17) and (18) imply (16).

Similarly (15) generalizes to

$$(19) \quad (x-a)\sum_0^k r_\nu(\xi) r_\nu(x) - (b-x)\sum_0^k s_\nu(\xi) s_\nu(x) = B_k p_{k+1}(x) q_k(\xi).$$

Indeed, denoting the left side by $\Psi(\xi, x)$ we find

$$\int_a^b \Psi(\xi, x)\pi_k(x) d\chi(x) = \pi_k(\xi) - \pi_k(\xi) = 0$$

hence

$$\Psi(\xi, x) = C_3(\xi) p_{k+1}(x).$$

Likewise

$$\int_a^b \Psi(\xi, x)(\xi - a)(b - \xi)\pi_{k-1}(\xi) d\chi(\xi)$$
$$= (x - a)(b - x)\pi_{k-1}(x) - (b - x)(x - a)\pi_{k-1}(x) = 0$$

and therefore

$$\Psi(\xi, x) = C_4(x) q_k(\xi).$$

Both results imply (19) and our identities (14), (15) are thereby established.

An improvement of our relations (11) and (13) is now described by

THEOREM 2. 1. *Let α_ν and β_ν be the zeros in increasing order of the polynomials $r_k(x)$ and $s_k(x)$ respectively. Between these zeros we have the inequalities*

$$(20) \qquad \beta_1 < \alpha_1 < \beta_2 < \alpha_2 < \ldots < \alpha_{k-1} < \beta_k < \alpha_k.$$

The real axis may therefore be decomposed into the following two sets of alternating intervals

$$(21) \qquad (-\infty, \beta_1],\ [\alpha_1, \beta_2], \ldots, [\alpha_{k-1}, \beta_k],\ [\alpha_k, +\infty),$$

$$(22) \qquad [\beta_1, \alpha_1],\ [\beta_2, \alpha_2], \ldots, [\beta_k, \alpha_k].$$

If $n = 2k$ then

$$(23) \quad \max_{f \in \Pi_n} f(z) = \begin{cases} \Phi_1(z) & \text{if } z \text{ is in one of the intervals (21).} \\ \Phi_2(z) & \text{if } z \text{ is in one of the intervals (22).} \end{cases}$$

2. *Let γ_ν and δ_ν be the zeros in increasing order of the polynomials $p_{k+1}(x)$ and $q_k(x)$, respectively. Between these zeros we have the inequalities*

$$(24) \qquad \gamma_1 < \delta_1 < \gamma_2 < \delta_2 < \ldots < \gamma_k < \delta_k < \gamma_{k+1}.$$

These points divide the axis in two sets of alternating intervals

$$(25) \qquad [\gamma_1, \delta_1], [\gamma_2, \delta_2], \ldots, [\gamma_k, \delta_k], [\gamma_{k+1}, +\infty)$$

$$(26) \qquad (-\infty, \gamma_1], [\delta_1, \gamma_2], \ldots, [\delta_k, \gamma_{k+1}].$$

If $n = 2k+1$ *then*

$$(27) \quad \max_{f \in \Pi_n} f(z) = \begin{cases} \Psi_1(z) & \text{if } z \text{ is in one of the intervals (25).} \\ \Psi_2(z) & \text{if } z \text{ is in one of the intervals (26).} \end{cases}$$

Indeed, if for the moment we take for granted the inequalities (20), then it is clear by (14) that α_ν, β_ν are all the zeros of $\Phi_1(z) - \Phi_2(z)$ and that all these zeros are simple. But then $\Phi_1(z) > \Phi_2(z)$ or $\Phi_1(z) < \Phi_2(z)$, depending on whether z is in the interior of an interval (21) or (22). Now (23) follows from (11). A similar argument establishes (27).

3. Proofs of the inequalities (20) and (24).

We shall use the known fact [3, pp. 46—47] that if c_1 and c_2 are real constants, $c_1 < c_2$, then the zeros of each of the polynomials

$$(28) \qquad q_k(x) - c_1 q_{k-1}(x), \; q_k(x) - c_2 q_{k-1}(x),$$

are real and simple, that they separate each other so that the least zero of the first polynomials is below the least zero of the second polynomial. Naturally, the zeros of the polynomials

$$(29) \qquad r_{k+1}(x) - c_1' r_k(x), \; r_{k+1}(x) - c_2' r_k(x), \quad (c_1' < c_2'),$$

enjoy similar properties.

We observe next that our polynomials satisfy the following two pairs of identities

$$(30) \qquad a_k s_k(x) = q_k(x) - c_1 q_{k-1}(x)$$

$$(30') \qquad b_k r_k(x) = q_k(x) - c_2 q_{k-1}(x), \qquad (c_1 < 0 < c_2),$$

$$(31) \qquad a_k' p_{k+1}(x) = r_{k+1}(x) - c_1' r_k(x),$$

$$(31') \quad b_k'(x-b) q_k(x) = r_{k+1}(x) - c_2' r_k(x), \qquad (c_1' < 0 < c_2'),$$

for appropriate constants c_i, c_i' having signs as indicated. It should be clear that the inequalities (20), (24) are easily established if we apply to the pairs of identities (30), (31) our introductory remark concerning the zeros of the polynomials (28) and (29).

Let us finally establish the identities (30)—(31′), for instance (30′). To this purpose we expand $r_k(x)$ in terms of the $q_\nu(x)$ in Fourier-fashion. If $\nu < k-1$ we find

$$\int_a^b r_k(x) q_\nu(x)(x-a)(b-x)d\chi(x) = 0$$

because the degrees of $q_\nu(x)(b - x)$ does not exceed $k - 1$. Therefore we indeed have an identity (30′). Now $-c_2$ has the same sign as

$$\begin{aligned}&\int_a^b r_k(x)q_{k-1}(x)(x-a)(b-x)d\chi(x)\\ &=\int_a^b r_k(x)\{q_{k-1}(x)(b-x)\}(x-a)d\chi(x)\\ &= -\int_a^b r_k(x)\{xq_{k-1}(x)\}(x - a)d\chi(x)\end{aligned}$$

which is negative. Therefore $c_2 > 0$. A similar argument proves (30) and that $c_1 < 0$. There is no need to reproduce the proofs of the remaining identities which use similar arguments.

4. Another extremum problem.

We consider the analogue of the previous problem for the interval $[0, +\infty)$. Let $d\chi(x)$ be again a measure not reducing to a finite number of point-masses. For each n we form the set Π_n of all polynomials of that degree satisfying the following relations:

$$f(x) \geqq 0 \text{ for } x \geqq 0, \int_0^\infty f(x)d\chi(x) = 1. \tag{32}$$

If z is arbitrary real, we may ask for the interval R_z representing the range of the values $f(z)$.

We can be brief. Let us denote by

$$\{p_\nu(x)\}, \ \{r_\nu(x)\}, \tag{33}$$

the orthonormal sets of polynomials associated with $d\chi(x)$ and $xd\chi(x)$ respectively, the highest coefficients of these polynomials being positive. We have then [cf. 3, p. 182]

$$f(x) = \begin{cases} |\sum_0^k u_\nu p_\nu(x)|^2 + x\,|\sum_0^{k-1} v_\nu r_\nu(x)|^2 & \text{if } n = 2k,\\ |\sum_0^k u_\nu p_\nu(x)|^2 + x\,|\sum_0^k v_\nu r_\nu(x)|^2 & \text{if } n = 2k+1,\end{cases} \tag{34}$$

where u_ν, v_ν are arbitrary *complex* numbers subject to the condition

$$\sum_0^k |u_\nu|^2 + \sum_0^k |v_\nu|^2 = 1, \quad (v_k = 0 \text{ if } n = 2k). \tag{35}$$

We introduce the abbreviations

$$\begin{aligned}&\Phi_1(x) = \sum_0^k (p_\nu(x))^2, \ \Phi_2(x) = x\sum_0^{k-1} (r_\nu(x))^2\\ &\Psi_1(x) = x\sum_0^k (r_\nu(x))^2, \ \Psi_2(x) = \Phi_1(x).\end{aligned} \tag{36}$$

We deal only with the case $z \geqq 0$ and only with the determination of $\max f(z)$. The result is:

$$(37)\qquad \max_{f\in\Pi_n} f(z) = \begin{cases} \max\,\{\Phi_1(z),\ \Phi_2(z)\} & \text{if } n = 2k, \\ \max\,\{\Psi_1(z),\ \Psi_2(z)\} & \text{if } n = 2k+1. \end{cases}$$

In order to distinguish between the quantities in the curly brackets, we employ the following identities:

$$(38)\qquad \sum_0^k (p_\nu(x))^2 - x\sum_0^{k-1}(r_\nu(x))^2 = A_k r_k(x) p_k(x),\ A_k > 0,$$

$$(38')\qquad \sum_0^k (p_\nu(x))^2 - x\sum_0^{k}(r_\nu(x))^2 = -B_k r_k(x) p_{k+1}(x),\ B_k > 0.$$

More generally we shall prove the "polarized" identites:

$$(39)\qquad \sum_0^k p_\nu(\xi)p_\nu(x) - x\sum_0^{k-1} r_\nu(\xi)r_\nu(x) = A_k r_k(\xi)p_k(x),$$

$$(39')\qquad \sum_0^k p_\nu(\xi)p_\nu(x) - x\sum_0^{k} r_\nu(\xi)r_\nu(x) = -B_k r_k(\xi)p_{k+1}(x).$$

For the proof of (39) we denote the left-hand side by $\Phi(\xi, x)$ and show as in the case of the finite interval:

$$\int_0^\infty \Phi(\xi, x)\pi_{k-1}(x)d\chi(x) = \pi_{k-1}(\xi) - \pi_{k-1}(\xi) = 0,$$

$$\int_0^\infty \Phi(\xi, x)\xi\pi_{k-1}\xi d\chi(\xi) = x\pi_{k-1}(x) - x\pi_{k-1}(x) = 0.$$

Similarly, denoting the left-hand side of (39') by $\Psi(\xi, x)$, we have

$$\int_0^\infty \Psi(\xi, x)\pi_k(x)d\chi(x) = \pi_k(\xi) - \pi_k(\xi) = 0,$$

$$\int_0^\infty \Psi(\xi, x)\xi\pi_{k-1}(\xi)d\chi(\xi) = x\pi_{k-1}(x) - x\pi_{k-1}(x) = 0.$$

This yields the assertions.

Further we point out [1, pp. 29—30] that

$$(40)\qquad x r_k(x) = \text{const}\left\{\frac{p_{k+1}(x)}{p_{k+1}(0)} - \frac{p_k(x)}{p_k(0)}\right\},$$

so that $r_k(x) = 0$ is equivalent with

$$p_{k+1}(x)/p_k(x) = p_{k+1}(0)/p_k(0) < 0,\ x > 0.$$

Thus denoting by α_ν, β_ν, γ_ν the zeros in increasing order of the polynomials $r_k(x)$, $p_k(x)$, $p_{k+1}(x)$ respectively, we have

$$(41)\quad \gamma_1 < \beta_1 < \alpha_1 < \gamma_2 < \beta_2 < \alpha_2 < \cdots < \gamma_k < \beta_k < \alpha_k < \gamma_{k+1}.$$

This furnishes

THEOREM 3 [2]). *We decompose the real axis into the following two sets of alternating intervals:*

$$(42)\quad (-\infty, \beta_1],\ [\alpha_1, \beta_2],\ [\alpha_2, \beta_3], \ldots, [\alpha_{k-1}, \beta_k],\ [\alpha_k, +\infty)$$

$$(43)\quad [\beta_1, \alpha_1],\ [\beta_2, \alpha_2], \ldots, [\beta_k, \alpha_k].$$

If $n = 2k$, *then*

$$(44)\quad \max_{f \in \Pi_n} f(z) = \begin{cases} \Phi_1(z) & \textit{if } z \textit{ is in one of the intervals } (42), \\ \Phi_2(z) & \textit{if } z \textit{ is in one of the intervals } (43). \end{cases}$$

Similarly we decompose the real axis as follows:

$$(45)\quad [\gamma_1, \alpha_1],\ [\gamma_2, \alpha_2], \ldots, [\gamma_k, \alpha_k],\ [\gamma_{k+1}, +\infty]$$

$$(46)\quad (-\infty, \gamma_1],\ [\alpha_1, \gamma_2],\ [\alpha_2, \gamma_3], \ldots, [\alpha_k, \gamma_{k+1}].$$

If $n = 2k + 1$, *then*

$$(47)\quad \max_{f \in \Pi_n} f(z) = \begin{cases} \Psi_1(z) & \textit{if } z \textit{ is in one of the intervals } (45), \\ \Psi_2(z) & \textit{if } z \textit{ is in one of the intervals } (46). \end{cases}$$

In conclusion we add a few words on the background of the problems of this note. The first paper devoted to such problems is due to F. Lukács [2]. Already Lukács was in possession of the two methods available for their solution: 1) The parametric representation of the polynomials, 2) Formulae of mechanical quadrature. In [2] Lukács uses the second method, while we have just used the first one. However, our results such as (23) were first derived by the second method. The evident continuity of the left-hand sides of (23) and (27) as functions of z, immediately furnishes as a byproduct the identities (14), (15), as was kindly pointed out to one of us by N. G. de Bruijn. Subsequently the

---

[2]) The problem discussed in this section is a limit case of our original extremum problem (§§ 1,3) if $a = 0$, $b = +\infty$. It clearly required the separate discussion which it received in the present § 4. However, the following remark is in order: *All polynomials, identities and results of* §§ 1, 2 *go formally over into the corresponding ones of* § 4 *if we set* $a = 0$ *and simply supress everywhere where it appears the factor* $b - x$.

Following this convention we see from (5) that the old polynomials defined by (4)

$$p_\nu(x),\ q_\nu(x),\ r_\nu(x),\ s_\nu(x)$$

now become

$$p_\nu(x),\ r_\nu(x),\ r_\nu(x),\ p_\nu(x)$$

respectively; also the Φ_i, Ψ_i of (9) turn into the similar symbols of (36). Moreover, the identities (14), (15) reduce to (38) and (38′) respectively. Finally Theorem 2 reduces to Theorem 3. However, these interesting formal relationships make our separate discussion of the case of the interval $[0, +\infty)$ no less necessary.

following soon became apparent: Once these identities are established directly (as we did), the parametric method proved to be a much more efficient way of establishing Theorems 2 and 3. In a forthcoming paper [1] L. Brickman discusses the problem of determining the domain D_z of variability of $f(z)$ ($f \in \Pi_n$), where z is a given complex (non-real) number. In this case it seems appropriate to turn once more to the method of mechanical quadratures.

University of Pennsylvania
and
Stanford University

REFERENCES

L. Brickman,

[1] A new generalization of a problem of F. Lukács. Compositio Mathematica, Volumen XIV, Fasc. 3.

F. Lukács,

[2] Verschärfung des ersten Mittelwertsatzes der Integralrechnung, Math. Zeitschrift 2 (1918), 295—305.

G. Szegö,

[3] Orthogonal Polynomials, Revised edition, New York, 1959.

(Oblatum 16-3-59).