On zeros of polynomials in best $L^p$-approximation and inserting mass points

Kenier Castillo¹ · Marisa de Souza Costa² · Fernando Rodrigo Rafaeli²

Accepted: 22 December 2020 / Published online: 27 January 2021
© Instituto de Matemática e Estatística da Universidade de São Paulo 2021

Abstract
The purpose of this note is to revive in $L^p$ spaces the original A. Markov ideas to study monotonicity of zeros of orthogonal polynomials.

Keywords Polynomials · Minimal $L^p$ norm · Zeros

Mathematics Subject Classification MSC 30C15

1 Introduction and main results

In History of Functional Analysis [6], p. 60], Dieudonné wrote: “The method of least squares of Legendre-Gauss had led Tchebychef to define a “best approximation” to a function...”. In contemporary language, the underlying problem considered by Chebyshev in the 1850’s can be rewritten in terms of the distance between points in

Communicated by José Alberto Cuminato.

This work was partially supported by the Centre for Mathematics of the University of Coimbra – UID/MAT/00324/2019, funded by the Portuguese Government through FCT/MEC and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020. FRR is supported by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAMEMIG) Demanda Universal under the grant APQ-03782-18, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under the grant 307267/2018-0.
metric spaces. Indeed, the problem of best approximation in a metric space, say $X$, by a fixed subset $Y$ consists in finding, for $f \in X$, an element $g_0 \in Y$ such that

$$d(f, g_0) = \inf_{g \in Y} d(f, g),$$

$d$ being the metric in $X$. The natural framework for treating these problems are the normed spaces and, of course, the metric is that induced by the norm, i.e., $d(f, g_0) := \|f - g_0\|$. Following Singer [14, p. 15], $\mathcal{Q}_Y(f)$ denotes the set of all elements of best approximation of $f$ by elements of $Y$. In general the characterization of semi-Chebyshev spaces is a pretty complex problem. However, the normal spaces with which we deal in this paper are suitable Banach spaces which possess in addition a uniform convexity of the norm (and $Y$ will be a Chebyshev space) and we consider the best $L^p$ approximation for the very specific function $x^{n+1}$.

Let $\mu$ be a positive Radon measure on a compact set $A \subseteq \mathbb{R}$. For $1 < p < \infty$, the space $L^p(\mu)$ denotes the set of all equivalent classes of $\mu$-measurable functions $f$ such that $|f|^p$ is $\mu$-summable, endowed with the usual vector operations and with the norm

$$\|f\|_p := \left( \int |f(x)|^p d\mu(x) \right)^{1/p}. \quad (1)$$

Set $X := L^p(\mu)$. By a well known result [4, Corollary, p.403], $X$ is uniformly convex. Define $\mathbb{N} := \{0, 1, \ldots \}$. Fix $n \in \mathbb{N}$ and set $K := \mathcal{P}_n$, $\mathcal{P}_n$ being the set of all real polynomials of degree at most $n$ regarded as a subspace of $X$. Since $K$ is finite dimensional, $K$ is a closed convex subspace of $X$. It is known that for any point $f \in X$, there is a unique point $g_0 \in \mathcal{Q}_K(f)$ (cf. [12, Theorem 8, p. 45]). The preceding affirmation thus guarantees the existence and uniqueness of $g_0 \in \mathcal{Q}_K(x^{n+1})$. By the characterization of elements of best approximation (cf. [14, Theorem 1.11]) $g_0 \in \mathcal{Q}_K(x^{n+1})$ if and only if

$$\int g(x)(x^{n+1} - g_0(x))^{|p-1|}\text{sgn}(x^{n+1} - g_0(x))d\mu(x) = 0 \quad (\forall g \in K). \quad (2)$$

Consider the (monic) polynomial $P_{n+1,p}(x) := x^{n+1} - g_0(x)$. As a consequence of (2), the minimum of the norm (1) taken over all (monic) real polynomials $P_{n+1}$ of degree $n+1$ is attained when $P_{n+1} := P_{n+1,p}$. By Fejér’s convex hull theorem (cf. [5, Theorem 10.2.2]), the zeros of $P_{n+1,p}$ all lie in the closure of the convex hull of supp($\mu$). Furthermore, all the zeros of $P_{n+1,p}$ are simple$^2$.

The central concern of this work is the following

---

1 If $\mathcal{Q}_Y(f) \neq \emptyset$, $Y$ is called semi-Chebyshev space. If Card($\mathcal{Q}_Y(f)$) = 1, $Y$ is called Chebyshev space or Haar space.

2 Suppose, contrary to our claim, that $x_0$ is a multiple zero. From (2) we have

$$\int \frac{P_{n+1,p}(x)}{(x-x_0)^2} |P_{n+1,p}(x)|^{p-1}\text{sgn}(P_{n+1,p}(x))d\mu(x) = \int \frac{|P_{n+1,p}(x)|^p}{|x-x_0|^2} d\mu(x) = 0,$$

a contradiction.
Question (Q) Let $\mu$ be a positive Radon measure on a compact set $A \subset \mathbb{R}$. Assume that $d\mu(x, t)$ has the form\(^3\)

$$
d\alpha(x, t) + j(t)\delta_y(t),$$

(3)

where $d\alpha(x, t) := \omega(x, t)\nu(x)$ ($\omega$ is a positive weight and $\nu$ is a positive Radon measure) and, $j(t) \in \mathbb{R}_+$ and $y(t) \in \mathbb{R}$ are continuous differentiable functions of $t \in U$, $U$ being an open interval on $\mathbb{R}$. Determine sufficient conditions in order for the zeros of the polynomial $P_{n+1, p}(x, t)$ ($2 \leq p < \infty$) to be strictly increasing functions of $t$.

For a technical reason related with the derivative of a convolution, we avoided the case $1 < p < 2$ in order to consider a general positive Radon measure. Moreover, for reason of economy of exposition, we avoided the case in which we have infinitely many mass points. Even though the reader has to proceed with caution in these cases, under natural additional assumptions, Theorem 1 below remains true, mutatis mutandis. Indeed, in [2], the reader can find a detailed study of the case $p = 2$ when we have infinitely many mass points, using the ideas originally presented in this work. We recall that this solves an open problem posed by Ismail at the end of the 1980’s within the framework of orthogonal polynomials (cf. [9, Problem 1] and [10, Problem 24.9.1]). When (3) has the form $\omega(x, t)dx$ and $p = 2$, Question (Q) was studied as early as 1886 by A. A. Markov [13, p. 178], in a work with many lights and some shadows (see, for instance, [1, Section 1] for some historical remarks). When (3) has the form $\omega(x, t)dx$ and $p = 2$, Question (Q) was posed as an exercise in Freud’s book [8, Problem 16, p. 133] (a proof of such result can be found in the more recent book by Ismail [10, Theorem 7.1.1]). When (3) has the form $\omega(x, t)dx$, $A := [-1, 1]$, and $1 \leq p \leq \infty$, Question (Q) was studied by Kroó and Peherstorfer [11]. When (3) has the form $\omega(x)dx + j\delta_y(t)$ and $p = 2$, Question (Q) was considered in [3, Theorem 2.2] through a combination of elementary facts. It is, therefore, natural that this last result be broadened to $L^p$ spaces. Not surprisingly, this can be easily achieved by using Markov’s original ideas.\(^4\) Our main result reads as follows:

Theorem 1 Assume the notation and conditions of Question (Q). Assume further the existence and continuity for each $x \in A$ and $t \in U$ of $(\partial \omega/\partial t)(x, t)$. Denote by $x_0(t), \ldots, x_n(t)$ the zeros of $P_{n+1, p}(x, t)$. Fix $k \in \{0, \ldots, n\}$ and set

$$
d_k(t) := \begin{cases} 
  y(t) - x_k(t) & \text{if } y(t) \neq x_k(t), \\
  1 & \text{if } y(t) = x_k(t).
\end{cases}
$$

\(^3\) Recall that the Dirac measure $\delta_y$ is a positive Radon measure whose support is the set $\{y\}$.

\(^4\) In his classical book [16, Footnote 31, p. 116], Szegő refers his proof of Markov’s theorem in the following terms: “This proof does not differ essentially from the original one by Markov, although the present arrangement is somewhat clearer.” Probably this assertion has avoided the attention of some mathematicians to Markov’s work. While it is true that in the framework of orthogonal polynomials Szegő’s argument becomes especially elegant, Markov’s approach works in a more general framework. Szegő’s approach is based on Gauss mechanical quadrature, which was an approach that Stieltjes suggested to handle the problem, see [15, Section 5, p. 391].
Define the function

\[ R_k(t) := \sum_{j=0}^{n} \frac{p - \delta_{jk}}{y(t) - x_j(t)}, \]

where the prime means that the sum is over all values \( j \) and \( t \) for which \( y(t) \neq x_j(t) \).

Then \( \frac{dx_k}{dt}(t) \) is strictly positive for those values of \( t \) such that

\[ \frac{1}{d_k(t)} \left( \frac{j'(t)}{j(t)} + y'(t)R_k(t) - \frac{1}{\omega(x_k(t), t)} \frac{\partial \omega}{\partial t}(x_k(t), t) \right) \geq 0, \tag{4} \]

and

\[ \frac{1}{\omega(x, t)} \frac{\partial \omega}{\partial t}(x, t) \tag{5} \]

is an increasing function of \( x \in A \), provided that at least the inequality (4) be strict or the function (5) be nonconstant on \( A \).

The next observations concern the cases studied in the literature for \( p = 2 \). As far as we know, these are the only ones that have been studied up to now. It is worth highlighting that such cases are the simplest consequences that can be derived from Theorem 1.

**Corollary 1** ² Let \( p = 2 \) and assume the notation and conditions of Theorem 1 under the constraint that \( \partial \mu(x, t) = \partial \alpha(x) + j\delta_{y(t)} \). Define the sets

\[ B_- := \{ t \in U \mid y(t) \in \text{Co}(A)^c \land y'(t) < 0 \}, \]
\[ B_+ := \{ t \in U \mid y(t) \in \text{Co}(A)^c \land y'(t) > 0 \}. \]

Then all the zeros of \( P_{n+1,2}(x, t) \) are strictly decreasing (respectively, increasing) functions of \( t \) on \( B_- \) (respectively, on \( B_+ \)).

**Proof** We only prove the result concerning to the set \( B_+ \); the rest follows in the same way. In this case, (4) reduces to

\[ \frac{y'(t)}{y(t) - x_k(t)} \left( \frac{1}{y(t) - x_0(t)} + \cdots + \frac{2}{y(t) - x_k(t)} + \cdots + \frac{1}{y(t) - x_n(t)} \right) > 0, \]

and (5) is equal zero.

---

² Corollary 1 for \( p = 2 \) was proved for the first time in [3, Theorem 2.2]. In order to have monotonicity of zero the location of the mass point outside \( \text{Co}(A) \) is quite natural.

³ \( A^c := \{ x \in \mathbb{R} \mid x \notin A \} \) and \( \text{Co}(A) \) denotes the convex hull of \( A \).
Corollary 2 Let \( p = 2 \) and assume the notation and conditions of Theorem 1 under the constraint that \( d\mu(x, t) = d\alpha(x) + j(t)d\gamma \). Define the sets
\[
C_- := \{ t \in U \mid j'(t) < 0 \}, \quad C_+ := \{ t \in U \mid j'(t) > 0 \}.
\]
If \( x_k(t) < y \) (respectively, \( x_k(t) > y \)) for each \( t \in U \), then \( x_k(t) \) is a strictly increasing (respectively, decreasing) function of \( t \) on \( C_+ \) (respectively, on \( C_- \)).

**Proof** As in Corollary 1, we only prove the result concerning to the set \( C_+ \) assuming \( x_k(t) < y \); the rest follows in the same way. In this case, (4) reduces to
\[
\frac{j'(t)}{j(t)} \frac{1}{y - x_k(t)} > 0,
\]
and (5) is equal zero.

The proof of Theorem 1 rests on two pillars: one is the characterization of elements of best approximation (2) and the other one is the implicit function theorem (cf. [7, Chapter III, Section 9]). Markov used the orthogonality relation that yields (2) when \( p = 2 \) (cf. [13, Equation 2]) together with the chain rule (cf. [13, Equation 5]), assuming that the zeros are implicitly defined as differentiable functions of the parameter. Kroó and Peherstorfer have also followed this approach in [11], using, in addition, the implicit function theorem to prove that the zeros are differentiable functions of the parameter. In some steps of our proof, the reader will be addressed to the corresponding step in Markov’s work.

2 Proof of Theorem 1

**Differentiability of the zeros:** Let \( P_{n+1}(x) := (x - x_0) \cdots (x - x_n) \), \( x_j \in \mathbb{R} \) \((j = 0, \ldots, n)\). (Note that the \( x_j \)'s do not depend on \( t \).) Define the map \( f := (f_0, \ldots, f_n) : \mathbb{R}^{n+1} \times U \to \mathbb{R}^{n+1} \), where we have set \( x := (x_0, \ldots, x_n) \) and
\[
f_k(x, t) := \int \frac{|P_{n+1}(x)|^p}{x - x_k} d\mu(x, t).
\]
For \( j \neq k \) one has
\[
\frac{\partial f_k}{\partial x_j}(x, t) = p \int \frac{1}{x - x_k} \frac{\partial P_{n+1}(x)}{\partial x_j} |P_{n+1}(x)|^{p-1} \text{sgn}(P_{n+1}(x)) d\mu(x, t);
\]
otherwise
\footnote{Cf. the denominator on the right-hand side of [13, Equation 5].}
\[
\frac{\partial f_k}{\partial x_k}(x, t) = \int \left| P_{n+1}(x) \right|^p \left| \frac{\partial}{\partial x_k} \frac{|x - x_k|^p}{x - x_k} \right| d\mu(x, t) \\
= (1 - p) \int \frac{|x - x_k|^p}{(x - x_k)^2} d\mu(x, t).
\]

(9)

Set \( x(t) := (x_0(t), \ldots, x_n(t)) \). Fix \( t_0 \in U \). From (7), (8) and (9), and using (2) we obtain

\[
f(x(t_0), t_0) = 0, \quad \frac{\partial f}{\partial x}(x(t_0), t_0) = \begin{pmatrix}
\frac{\partial f_0}{\partial x_0}(x(t_0), t_0) \\
\vdots \\
\frac{\partial f_n}{\partial x_n}(x(t_0), t_0)
\end{pmatrix},
\]

with

\[
\frac{\partial f_j}{\partial x_j}(x(t_0), t_0) \neq 0 \quad \text{for all} \quad j = 0, 1, \ldots, n.
\]

According to the implicit function theorem, under these conditions the equation \( f(s, t) = 0 \) has a solution \( s = x(t) \) in a neighborhood of \((x(t_0), t_0)\) which is a differentiable function on \( t \).

Expression for the derivative of the zeros: In view of the above result\(^9\),

\[
\frac{dx_k}{dt}(t) = -\frac{\frac{\partial f_k}{\partial x_k}(x(t), t)}{\frac{\partial f_k}{\partial x_k}(x(t), t)}.
\]

(10)

We see at once that

\[
\frac{\partial f_k}{\partial t}(x(t), t) = \int \left| P_{n+1, \rho}(x, t) \right|^p \frac{\partial \omega}{\partial t}(x, t) \, dv(x) \\
+ \left( j'(t) + j(t) y'(t) R_k(t) \right) \frac{|P_{n+1, \rho}(y(t), t)|^p}{y(t) - x_k(t)).
\]

(11)

Clearly\(^{10}\)

\[
\frac{1}{\omega(x_k(t), t)} \frac{\partial \omega}{\partial t}(x_k(t), t) \int \left| P_{n+1, \rho}(x, t) \right|^p \frac{1}{x - x_k(t)} d\mu(x, t) = 0.
\]

Subtracting this from the left-hand side of (11) yields\(^{11}\)

\(^9\) Cf. the left-hand side of [13, Equation 5].

\(^{10}\) Cf. [13, p. 179].

\(^{11}\) Cf. the numerator on the right-hand side of [13, Equation 5].
\[
\frac{\partial f_k}{\partial t}(x(t), t) = \int \frac{|P_{n+1,p}(x, t)|^p}{x - x_k(t)} dx(t) \tag{12}
\]

\[
\left( \frac{1}{\omega(x, t)} \frac{\partial \omega}{\partial t}(x, t) - \frac{1}{\omega(x_k(t), t)} \frac{\partial \omega}{\partial t}(x_k(t), t) \right) \omega(x, t) dv(x) + \left( f'(t) + j(t) y'(t) R_k(t) - \frac{j(t)}{\omega(x_k(t), t)} \frac{\partial \omega}{\partial t}(x_k(t), t) \right) \frac{|P_{n+1,p}(y(t), t)|^p}{y(t) - x_k(t)}. \tag{13}
\]

It only remains to note that\(^\text{12}\)

\[
\frac{1}{x - x_k(t)} \left( \frac{1}{\omega(x, t)} \frac{\partial \omega}{\partial t}(x, t) - \frac{1}{\omega(x_k(t), t)} \frac{\partial \omega}{\partial t}(x_k(t), t) \right) \geq 0.
\]

Thus

\[
\text{sgn} \left( \frac{dx_k}{dt}(t) \right) = \text{sgn} \left( \frac{df_k}{dt}(x(t), t) \right),
\]

and the desired result follows from (1213).

References

1. Castillo, K.: On monotonicity of zeros of paraorthogonal polynomials on the unit circle. Linear Algebra Appl. 580, 474–490 (2019)
2. Castillo, K., Costa, M.S., Rafaeli, F.R.: On Markov’s theorem on zeros of orthogonal polynomials revisited. Appl. Math. Comput. 339, 390–397 (2018)
3. Castillo, K., Rafaeli, F.R.: On the discrete extension of Markov’s theorem on monotonicity of zeros. Electron. Trans. Numer. Anal. 44, 271–280 (2015)
4. Clarkson, J.A.: Uniformly convex spaces. Trans. Am. Math. Soc. 40, 415–420 (1936)
5. Davis, P.J.: Interpolation and Approximation. Republication, with Minor Corrections, of the: Original, with a New Preface and Bibliography, p. 1975. Dover Publications Inc, New York (1963)
6. History of Functional Analysis, North-Holland Mathematics Studies, 49. Notas de Matemática [Mathematical Notes], vol. 77. North-Holland Publishing Co., Amsterdam (1981)
7. Lima, E. L.: Curso de Análise, Vol. 2. (2nd ed.) (Portuguese) [Course in Analysis, Vol. 2], Projeto Euclides [Euclids Project], 13. Instituto de Matemática Pura e Aplicada, Rio de Janeiro (1981)
8. Freud, G.: Orthogonal Polynomials. Pergamon Press, Oxford (1971)
9. Ismail, M. E.: Monotonicity of zeros of orthogonal polynomials. In: q-Series and Partitions (Minneapolis, MN, 1988), volume 18 of IMA Vol. Math. Appl., pp. 177-190. Springer, New York (1988)
10. Ismail, M.E.H.: Classical and Quantum Orthogonal Polynomials in One Variable. Encyclopedia of Mathematics and Its Applications, vol. 98. Cambridge University Press, Cambridge (2005)
11. Kroó, A., Pehlerstorfer, F.: On the zeros of polynomials of minimal \(L_p\)-norm. Proc. Am. Math. Soc. 101, 652–656 (1987)
12. Lax, P.: Functional Analysis, Pure and Applied Mathematics. Wiley, New York (2002)

\(^{12}\) Cf. [13, p. 179].
13. Markov, A.: Sur les racines de certaines équations (second note). Math. Ann. 27, 177–182 (1886)
14. Singer, I.: Best approximation in normed linear spaces by elements of linear subspaces, Translated from the Romanian by Radu Georgescu. Die Grundlehren der mathematischen Wissenschaften, Band 171 Publishing House of the Academy of the Socialist Republic of Romania, Bucharest. Springer, New York (1970)
15. Stieltjes, T.J.: Sur les racines de l’équation $X_n = 0$. Acta Math. 9, 385–400 (1887)
16. Szegő, G.: Orthogonal polynomials, volume 23. Amer. Math. Soc. Coll. Publ., Amer. Math. Soc., Providence, R. I., 4th edition, 1975 edition (1939)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.