Filomat 38:20 (2024), 7305–7316
https://doi.org/10.2298/FIL2420305B

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Polynomial approximation of L^2-functions

Nour Elhouda Boucherchem[a,*], **Anis Rezgui**[b,*]

[a] *Mathematics Department, ESST Hammam Sousse, University of Sousse, Tunisia*
[b] *Mathematics Department, INSAT, Carthage University, Tunisia*

Abstract. Let μ be a given probability measure supported by a compact subset $[a,b] \subset \mathbb{R}$. Given a function f element of $L^2\Big([a,b], d\mu\Big)$, we proved, under some integrability conditions, that a continuous version of f can be pointwisely and uniformly approximated by a sequence of polynomial functions. More precisely by a partial-sum of orthogonal polynomials in $L^2\Big([a,b], d\mu\Big)$. As an application, we have used the obtained approximation Theorem to set up a polynomial interpolation algorithm of L^2-functions. The derived interpolation algorithm has been implemented and compared to standard ones, such as the spline-cubic one.

1. Introduction

In 1885 Weierstrass has proved a corner stone result in approximation theory, that is, every continuous function on a compact subset, can be pointwisely and uniformly approximated by a polynomial function, see [12]. This result has been generalized to an abstract functional analysis framework by M. H. Stone in 1937, see [18], no doubt that this both results has constituted the starting point of approximation theory, see [6][19] for a more comprehensive study.

Since then, numerous research studies have been conducted on the subject of approximating functions by polynomial functions, serving various purposes such as mathematical problems, computer science, engineering and environmental sciences, etc ... Here is a selection of papers that highlight the range of investigations carried out, see [16][21][9][10][5].

Although continuity condition is quite essential to approximate functions by polynomials, in practical situations we are usually confronted with a design of points without any idea about the minimal regularity that should satisfy the generating function of that design. In general, we impose at least continuity, actually without justification, just to be able to run polynomial approximation algorithms.

Let μ be a probability measure on a compact subset of $\mathbb{R}$, $[a,b] \subset \mathbb{R}$, which we suppose to be absolutely continuous with respect to the Lebesgue measure i.e $d\mu(x) = \omega(x)dx$ for a given non-negative Borel function ω. It is known that the family of monomials $\{x^n : n \geq 0\}$ is total in the Hilbert space $L^2_\mu := L^2\Big([a,b], d\mu\Big)$,

2020 *Mathematics Subject Classification*. 41A05, 46C07, 46C15 and 30D15.
Keywords. Probability measures, Approximation of functions, Polynomial interpolation.
Received: 12 November 2023; Revised: 24 December 2023; Accepted: 01 January 2024
Communicated by Miodrag Spalević
* Corresponding author: Nour Elhouda Boucherchem
Email addresses: nourbou08@gmail.com (Nour Elhouda Boucherchem), anis.rezguii@gmail.com (Anis Rezgui)

see for instance [4], this implies existence of an orthonormal basis $\{e_n : n \geq 0\}$, such that every function φ element of L^2_μ can be expressed as follows

$$\varphi = \sum_{n=0}^{+\infty} \varphi^{(n)} e_n, \tag{1}$$

where the convergence is in L^2_μ and where $\varphi^{(n)} = \langle \varphi, e_n \rangle_2$, with $\langle \cdots, \cdots \rangle_2$ the scalar product in L^2_μ. For each $N \geq 0$ denote the truncated sum up to N by

$$\varphi_N(x) = \sum_{n=0}^{N} \varphi^{(n)} e_n(x)\cdot \tag{2}$$

Then the equation (1) can also be expressed as follows

$$\varphi_N \xrightarrow[N\to+\infty]{L^2_\mu} \varphi. \tag{3}$$

Our main concern, in the current paper, is to study the necessary conditions under which the L^2-convergence in (3) could be turned into pointwise and uniform convergence ? And as a result, any given L^2-function could hence be approximated pointwisely by a polynomial.

The paper is organized as follows, in the next section the main approximation result is stated as well as the essential steps to achieve it. In a third section the polynomial approximation result is converted into a polynomial interpolation algorithm, that is implemented and tested on many significant cases. The fourth section is just an appendix in which the proof of Theorem 2.2 is detailed since it is long and technical.

2. Pointwise approximation result

In this section the main approximation result Theorem 2.6 will be stated and proved, before let us introduce some notations and present some examples.

The probability μ is still as defined above i.e a given absolutely continuous probability measure supported by a compact subset $[a, b] \subsetneq \mathbb{R}$, $d\mu = \omega dx$ where ω is a non-negative Borel function. For all $n \in \mathbb{N}$ denote by $\mu_n = \int_a^b x^n d\mu(x)$ the n^{th} order moment of μ. Let $\{e_n : n \in \mathbb{N}\}$ be the orthonormal basis of L^2_μ obtained by the Gram-Schmidt procedure applied on the total subset $\{x^n : n \in \mathbb{N}\}$. It is shown in [17] that for all $n \geq 1$ the polynomial function e_n can be expressed in terms of moments as follows

$$e_n(x) = \frac{1}{\sqrt{d_n d_{n-1}}} \begin{vmatrix} \mu_0 & \cdots & \mu_{n-1} & 1 \\ \mu_1 & \cdots & \mu_n & x \\ \vdots & & & \\ \mu_n & \cdots & \mu_{2n-1} & x^n \end{vmatrix}, \tag{4}$$

where d_n is the determinant of the matrix D_n

$$d_n = |D_n| \text{ with } D_n = \begin{pmatrix} \mu_0 & \cdots & \mu_n \\ \mu_1 & \cdots & \mu_{n+1} \\ \vdots & & \\ \mu_n & \cdots & \mu_{2n} \end{pmatrix}. \tag{5}$$

This means in particular that any polynomial e_n is exactly of degree n and that for each $n \geq 1$

$$e_n(x) = \kappa_n x^n + \text{"lower order"}, \tag{6}$$

where the highest order coefficient satisfies

$$0 < \kappa_n = \sqrt{\frac{d_{n-1}}{d_n}}, \tag{7}$$

and $e_0 = \kappa_0 = 1$.
Next is a presentation of some known examples of probability measures for which the orthonormal polynomials are known and where we can compute explicitly κ_n, for more details we refer to [3].

Examples 2.1.

1. *Let μ be the uniform probability measure on $[-1,1]$ and $\{P_n(x)\}_{n\in\mathbb{N}}$ the family of Legendre polynomials. Then the polynomials family $\left\{\sqrt{(2n+1)}P_n(x)\right\}_{n\in\mathbb{N}}$ constitutes an orthonormal basis for the Hilbert space $L^2\left([-1,1],\frac{dx}{2}\right)$. And for every $n\in\mathbb{N}$*

$$\kappa_n = \sqrt{2n+1}\frac{(2n)!}{2^n(n!)^2}.$$

2. *Let μ be the arcsine probability defined on $[-1,1]$ by its density*

$$\omega(x) = \frac{1}{\pi\sqrt{1-x^2}},$$

and $\{T_n\}_{n\in\mathbb{N}}$ be the Chebyshev polynomials. Then the family $\left\{\sqrt{2}T_n(x) : n\in\mathbb{N}\right\}$ constitutes an orthonormal basis for the Hilbert space $L^2\left([-1,1],\frac{1}{\pi\sqrt{1-x^2}}dx\right)$. And for every $n\in\mathbb{N}$

$$\kappa_n = \begin{cases} 1 & \text{if } n = 0, \\ 2^{n-\frac{1}{2}} & \text{if } n \neq 0. \end{cases}$$

3. *For $\alpha > \frac{1}{2}$, we suppose μ to be the Gamma probability measure defined on $[-1,1]$ by its density function*

$$\omega(x) = \frac{\Gamma(\alpha+1)}{\sqrt{\pi}\Gamma(\alpha+\frac{1}{2})}(1-x^2)^{\alpha-\frac{1}{2}},$$

and we set $\left\{C_n^{(\alpha)}\right\}_{n\in\mathbb{N}}$ to be the so-called Gegenbauer polynomials. Then the polynomials family

$$\left\{\sqrt{\frac{n!(n+\alpha)\Gamma(\alpha)\Gamma(\alpha+\frac{1}{2})}{\alpha 2^{1-2\alpha}\sqrt{\pi}\Gamma(n+2\alpha)}}C_n^{(\alpha)}(x) : n\in\mathbb{N}\right\}$$

constitutes an orthonormal basis for the Hilbert space $L^2([-1,1],d\mu)$ and for every $n\in\mathbb{N}$

$$\kappa_n = (\alpha)_n 2^{n+\alpha-\frac{1}{2}}\sqrt{\frac{(n+\alpha)\Gamma(\alpha)\Gamma(\alpha+\frac{1}{2})}{\alpha n!\sqrt{\pi}\Gamma(n+2\alpha)}},$$

where $(\alpha)_n = \alpha(\alpha+1)\dots(\alpha+n-1)$ is the Pochhammer symbol.

In the next Theorem 2.2 it is shown that for all $x\in[a,b]$, the sequence $\{|e_n(x)|\}_{n\in\mathbb{N}}$ is not small enough, when n grows to infinity. And so the question of investigating pointwise convergence of the L^2-summation in (1) is actually non-trivial.

Theorem 2.2. *For all $x \in [a, b]$ we have*

$$\sum_{n=0}^{+\infty} |e_n(x)|^2 = +\infty. \tag{8}$$

Remark 2.3. *The proof of Theorem 2.2 is quite technical and requires some Lemmas, we postponed it to Appendix in Section 4.*

The probability measure μ is said to satisfy the Szegö's condition if

$$\int_a^b \frac{\log \omega(x)}{\sqrt{(x-a)(b-x)}} dx > -\infty \,. \tag{9}$$

Next are some examples which could be found for instance in [22].

Examples 2.4.

1. *If we consider the uniform probability measure on $[a, b]$, we have*

$$\int_a^b \frac{-\log(b-a)}{\sqrt{(x-a)(b-x)}} dx = -\pi \log(b-a).$$

2. *If we consider the family of Beta-probability measures defined on $[0, 1]$ for $\alpha, \beta > -1$ by*

$$d\mu(x) = \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)} x^{\alpha}(1-x)^{\beta} dx.$$

It is straightforward to obtain

$$\int_0^1 \frac{\log \omega(x)}{\sqrt{x(1-x)}} dx = -\pi \left(2(\alpha+\beta)\log 2 + \log\left(\frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)} \right) \right).$$

3. *If we consider the arcsine probability measure on $[-1, 1]$, we have*

$$\int_{-1}^1 \frac{-\log\left(\pi\sqrt{1-x^2}\right)}{\sqrt{1-x^2}} dx = \pi \left(\log 2 - \log \pi \right).$$

The next Lemma 2.5 gives a uniform bounds of the polynomial functions $e_n(\cdot)$ for all n$\in \mathbb{N}$, this Lemma is crucial for the sequel.

Lemma 2.5. *Suppose that the probability measure μ satisfies the Szegö's condition (9), then there exist $C, A > 0$ such that*

$$|e_n(x)| \leq CA^n, \tag{10}$$

for all $n \in \mathbb{N}$ and all $x \in [a, b]$.

Proof of Lemma 2.5
Since the probability measure μ is compactly supported it is known, see for instance [7], that for all $n \in \mathbb{N}$ the polynomial $e_n(\cdot)$ has exactly n zeros $x_1^{(n)}, \cdots, x_n^{(n)}$ that are all in $]a, b[$. Then for all $x \in [a, b]$ and for all $n \in \mathbb{N}$ we have

$$\begin{aligned} |e_n(x)| &= \kappa_n \prod_{k=1}^{n} \left| x - x_k^{(n)} \right| \\ &< \kappa_n \prod_{k=1}^{n} (b-a) \\ &< \kappa_n (b-a)^n . \end{aligned}$$

By using Lemma 4.5, see the appendix below, we finish the proof. □

For each integer $p \in \mathbb{N}$ the following sequence of norms is defined on L^2_μ

$$\|\varphi\|_p^2 = \sum_{n=0}^{+\infty} 2^{np} |\varphi^{(n)}|^2. \tag{11}$$

The sequence of norms $\left\{\| \ \|_p : \ p \in \mathbb{N}\right\}$ is increasing w.r.t $p \in \mathbb{N}$

$$\| \ \|_{p+1} \geq \| \ \|_p \geq \cdots \geq \| \ \|_0 = \| \ \|_{L^2_\mu},$$

and then it defines a sequence of decreasing sub-spaces of L^2_μ, for all $p \in \mathbb{N}$

$$\mathbb{H}_p := \left\{ \varphi \in L^2_\mu : \ \|\varphi\|_p < \infty \right\}. \tag{12}$$

By using a usual completion process a decreasing chain of Hilbert sub-spaces of $L^2\Big([a, b], \mu\Big)$ is obtained

$$\cdots \subsetneq \mathbb{H}_{p+1} \subsetneq \mathbb{H}_p \subsetneq \cdots \subsetneq \mathbb{H}_0 = L^2([a, b], \mu).$$

The main result of the current paper, is now ready to be stated

Theorem 2.6.
Suppose that the probability measure μ satisfies the Szegö's condition (9). Then there exists a positive integer $p > 0$ such that

1. *all functions elements of $\mathbb{H}_p$ admit a continuous version,*
2. *for φ a given function element of $\mathbb{H}_p$, its associated sequence of polynomial functions $\{\varphi_N : \ N \in \mathbb{N}\}$ defined in (2) convergences pointwisely and uniformly to φ, on $[a, b]$.*
3. *The following error control holds*

$$\sup_{x \in [a,b]} |\varphi(x) - \varphi_N(x)| \leq C B^N \|\varphi\|_p. \tag{13}$$

for all functions $\varphi \in \mathbb{H}_p$ and for some constants $C > 0$ and $0 < B < 1$.

Proof of Theorem 2.6
Let μ be a probability measure with a compact support and that satisfies the Szegö's condition, for all $x \in [a,b]$ and a given $p > 0$we have

$$\begin{aligned}\sum_{n\geq 0}\left|\varphi^{(n)}e_n(x)\right| &= \sum_{n\geq 0}2^{\frac{np}{2}}\left|\varphi^{(n)}\right|2^{\frac{-np}{2}}\left|e_n(x)\right| \\ &\leq \|\varphi\|_p\left[\sum_{n\geq 0}2^{-np}\left|e_n(x)\right|^2\right]^{1/2}.\end{aligned}$$

By using Lemma 2.5 we deduce that for all $x \in [a,b]$, the series $\sum_{n\geq 0} 2^{-np}|e_n(x)|^2$ converges for some $p > 0$.

1. Let φ be a function element of $\mathbb{H}_p$ for $p > 0$ given above, then the function defined as the uniform limit on $[a,b]$

$$\widetilde{\varphi}(x) = \sum_{n\geq 0}\varphi^{(n)}e_n(x),$$

is a continuous version of φ.

2. For φ element of $\mathbb{H}_p$, and φ_N the polynomial function for some $N > 0$

$$\varphi_N(x) = \sum_{n=0}^{N}\varphi^{(n)}e_n(x),$$

we have

$$\begin{aligned}|\varphi(x) - \varphi_N(x)| &\leq \sum_{n=N+1}^{+\infty}\left|\varphi^{(n)}e_n(x)\right| \\ &\leq \|\varphi\|_p\left[\sum_{n\geq N+1}2^{-np}\left|e_n(x)\right|^2\right]^{1/2} \\ &\leq \|\varphi\|_p\frac{C}{\sqrt{1-2^{-p}A^2}}\left[2^{-p/2}A\right]^{N+1},\end{aligned}$$

where C and A are the constants given by Lemma 2.5. This proves 2. and 3. and then finishes the proof of Theorem 2.6.

□

3. Numerical illustration

In this section Theorem 2.6 is used to work out a polynomial interpolation algorithm to recover a given real function that belongs to the space L^2_μ for μ a compactly supported probability measure that satisfies the Szegö's condition (9).

Let $(x_0, y_0), \ldots, (x_N, y_N)$ be a given design of points, with $N > 0$ a positive integer. Suppose that there exists φ an L^2_μ function that belongs to $\mathbb{H}_p$, for $p > 0$ given in Theorem 2.6, and suppose that $y_0 = \varphi(x_0), \ldots, y_N = \varphi(x_N)$. Since the function φ can be pointwisely approximated by the sequence of polynomial functions $\{\varphi_n : n \in \mathbb{N}\}$, if we suppose that N is large enough we can consider that $\varphi_N = \sum_{n=0}^{N} \varphi^{(n)} e_n$ approximates φ at the knots $\{x_0, \ldots, x_N\}$ i.e.

$$\begin{bmatrix} \varphi_N(x_0) \\ \vdots \\ \varphi_N(x_N) \end{bmatrix} \simeq \begin{bmatrix} \varphi(x_0) \\ \vdots \\ \varphi(x_N) \end{bmatrix},$$

which leads to the following system of linear equations

$$\begin{bmatrix} e_0(x_0) & \cdots & e_N(x_0) \\ \vdots & & \vdots \\ e_0(x_N) & \cdots & e_N(x_N) \end{bmatrix} \begin{bmatrix} \varphi^{(0)} \\ \vdots \\ \varphi^{(N)} \end{bmatrix} \simeq \begin{bmatrix} \varphi(x_0) \\ \vdots \\ \varphi(x_N) \end{bmatrix}. \tag{14}$$

By solving it, we derive the coefficients $\varphi^{(0)}, \cdots, \varphi^{(N)}$ and so we obtain the interpolation polynomial function

$$\varphi_N(x) = \sum_{n=0}^{N} \varphi^{(n)} e_n(x).$$

Remark 3.1. *Note that the accuracy of approximation depends on the initial design's size N, and doesn't depend neither on the distances between successive design's points nor on function's regularity*

$$|\varphi(x) - \varphi_N(x)| \leq CB^N \|\varphi\|_p. \tag{15}$$

However standard polynomial interpolation algorithms does, see for instance [11]

$$|\varphi(x) - p_N(x)| \leq \frac{\left(\max_{1\leq k\leq N} |x_k - x_{k-1}|\right)^{N+1}}{(N+1)!} \sup_{[a,b]} \left| \frac{d^{N+1}\varphi}{du^{N+1}}(u) \right|, \tag{16}$$

where p_N denotes the Taylor expansion polynomial interpolation of order N associated to φ.

Next we present three examples of probabilities and of functions to be interpolated. We consider a design's size $N = 5,\ 10,\ 15$. And then we compute for each N the quadratic error between the given function and the interpolation points for the well-known "cubic spline" algorithm and the one proposed in this paper.

Example 1. *Let μ be the arcsine probability defined on $[-1, 1]$ by its density function $\omega(x) = \dfrac{1}{\pi\sqrt{1-x^2}}$ and let $g(x) = \Gamma(x^2 + 1)$ be the function to be interpolated. In figure 1 we have a graphical illustration for $N = 5$.*

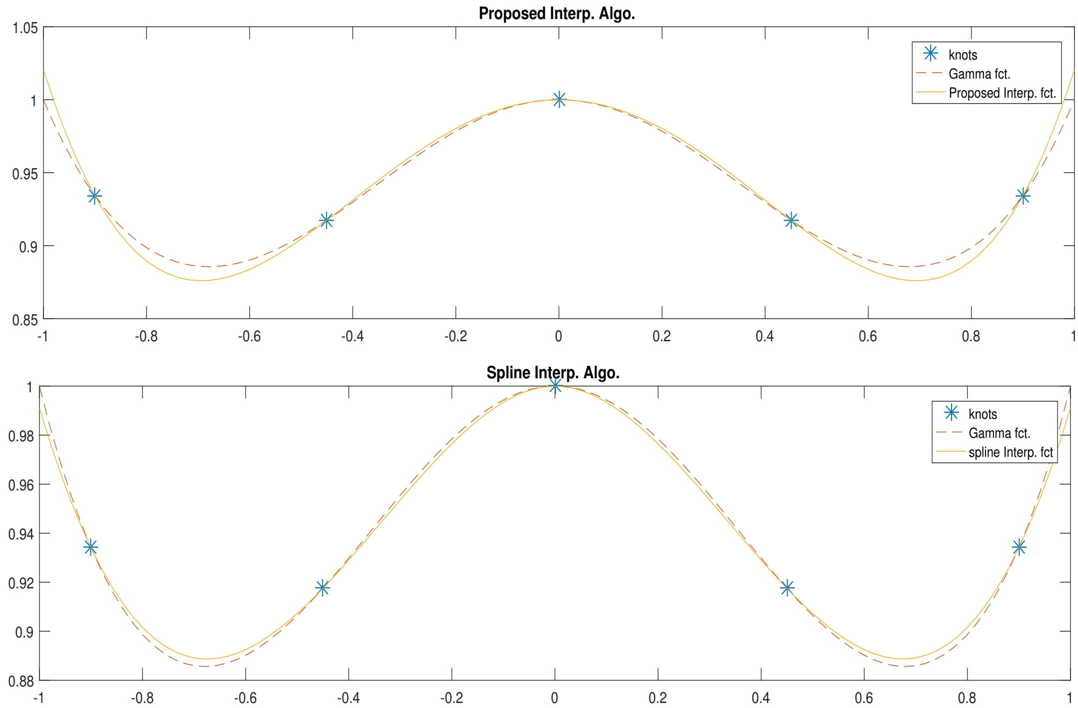

Figure 1: Proposed Interp. Algo. vs Spline Algo., $g(x) = \Gamma(x^2 + 1)$.

The table 1 records errors of the two algorithms for different values of $N = 5, 10$ *and* 15*:*

N	Proposed Algo. error	Spline error
5	0.0063	0.0025
10	0.0014	0.0007
15	0.3259e-003	0.4081e-003

Table 1: Proposed Algo. error vs Spline error of $g(x)$.

Remark 3.2. *Note that in example 1, we expect that the "Proposed algorithm" is more accurate for N large since the interpolated function is of exponential growth and so polynomial of order N is closer than cubic spline, this can be seen in table 1.*

Example 2. *Suppose that* μ *is the probability measure on* $[-1, 1]$ *defined by its density function* $\omega(x) = \frac{3}{4}(1 - x^2)$*. Let* $h(x) = 2x^4 + 4x^3 + x^2 - 5x + 3$ *be the function to be interpolated. The figure 2 is the graphical illustration for* $N = 5$*.*

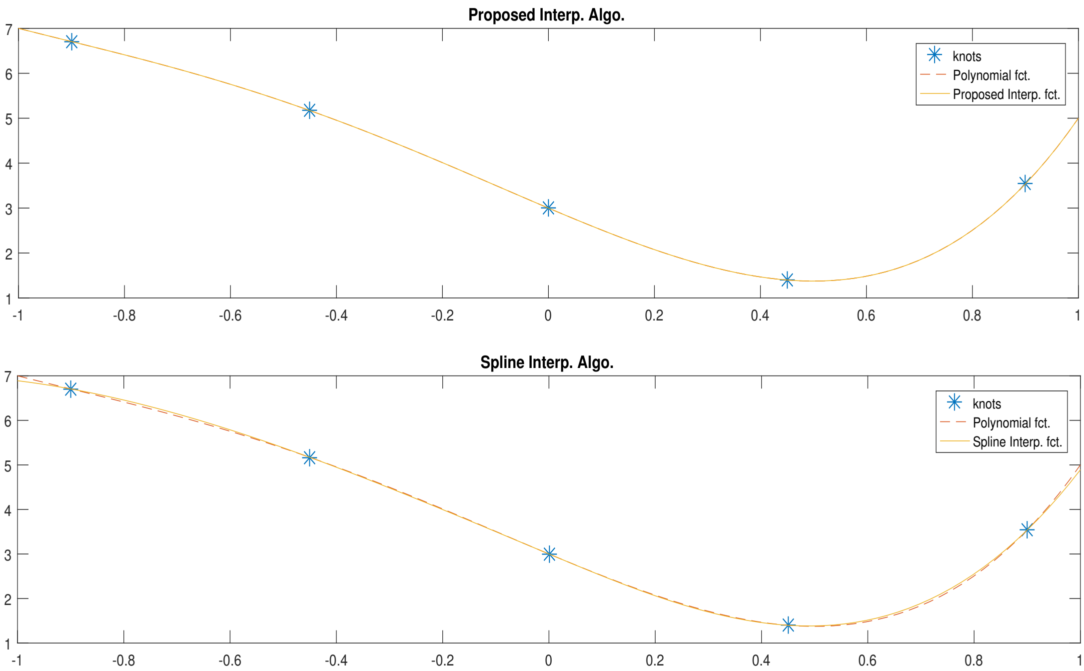

Figure 2: Proposed Interp. Algo. vs Spline Algo., $h(x) = 2x^4 + 4x^3 + x^2 - 5x + 3$.

In the table 2 we have recorded errors of the two algorithms for different values of $N = 5, 10$ *and* 15*:*

N	Proposed Algo. error	Spline error
5	8.9012e-015	325.3404e-003
10	41.7627e-015	29.8148e-003
15	696.5221e-015	10.7723e-003

Table 2: Proposed Algo. error vs Spline error of $h(x)$.

Remark 3.3. *In Example 2 the interpolated function is a polynomial of order 4 and the probability density is also a polynomial of order 2. The proposed algorithm shows better accuracy than the "cubic spline", for all N.*

Example 3. *Let* μ *be the uniform probability measure on* $[-1, 1]$ *and that the interpolated function* $\varphi(x) = \cos(3x)$*. The next Figure 3 shows a graphical illustration for* $N = 5$*.*

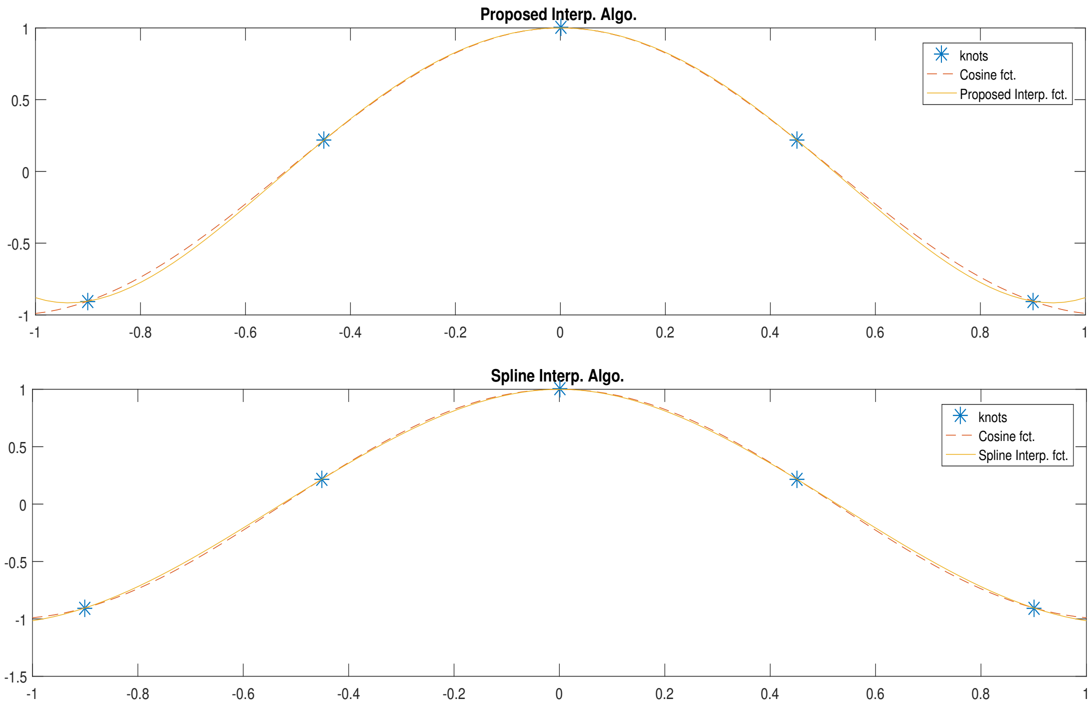

Figure 3: Proposed Interp. Algo. vs Spline Algo., $f(x) = \cos(3x)$.

In the table 3 we have recorded errors of the two algorithms for different values of $N = 5, 10$ *and* 15:

N	Proposed Algo. error	Spline error
5	0.0277	0.0145
10	0.000156	0.00277
15	0.78e-008	0.00131

Table 3: Proposed Algo. error vs Spline error of $f(x)$.

Remark 3.4. *In Example 3 the interpolated function is bounded but of polynomial type of infinite order, the proposed algorithm's accuracy is clearly better than the cubic spline one, when N increases.*

4. Appendix: Proof of Theorem 2.2

As mentioned above the current appendix is dedicated to the proof of Theorem 2.2. Next there are five needed lemmas, that deal about the growth of the highest order coefficient of the associated polynomial $e_n(\cdot)$, κ_n uniformly w.r.t $x \in [a, b]$. Actually the proof of these five lemmas can be found in the very self-contained paper [1] but also more recently in [7].

Lemma 4.1. *The following inequality holds for all* $n \in \mathbb{N}$

$$1 \leq \mu_{2n} \left[\kappa_n\right]^2 . \quad (17)$$

For $n \in \mathbb{N}$, denote by $a_n = \dfrac{\kappa_n}{\kappa_{n+1}}$ and by $b_n = \int_a^b x e_n^2(x) d\mu(x)$, and convince that $e_{-1} \equiv 0$ and $e_0 \equiv 1$, the following Lemma has been proved.

Lemma 4.2. *For all $x \in [a, b]$, the sequence of polynomials $\{e_n(x)\}_{n \geq -1}$ is solution of the following recursive functional equation*

$$\begin{cases} u_{-1}(x) = 0 \text{ and } u_0(x) = 1 \\ x u_n(x) = a_n u_{n+1}(x) + b_n u_n(x) + a_{n-1} u_{n-1}(x), \text{ for all } n \geq 1. \end{cases} \tag{18}$$

Let $\{q_n(\cdot)\}_{n \geq -1}$ be the solution of the recursive equation (18) with the initial conditions $u_{-1}(\cdot) = -1$ and $u_0(\cdot) = 0$, then the next Lemma is proved.

Lemma 4.3. *For all integers $n \in \mathbb{N}$*

$$a_n \left[q_{n+1}(\cdot) e_n(\cdot) - q_n(\cdot) e_{n+1}(\cdot) \right] = 1. \tag{19}$$

The next Lemma derives straightforwardly from Lemma 4.3 and Cauchy–Schwarz inequality.

Lemma 4.4. *For every integer $n \in \mathbb{N}$ and for all $x \in [a, b]$ we have*

$$\sum_{k=0}^{n} \frac{1}{a_k} \leq 2 \left[\sum_{k=0}^{n+1} |e_k(x)|^2 \right]^{1/2} \left[\sum_{k=0}^{n+1} |q_k(x)|^2 \right]^{1/2}. \tag{20}$$

While the Szegö's condition is unnecessary for the proof of the above four Lemmas, it is essential for the next Lemma.

Lemma 4.5. *If the probability measure satisfies moreover the Szegö's condition (9, then*

$$\limsup_n \kappa_n^{1/n} < \infty. \tag{21}$$

The proof of Theorem 2.2 is now ready to be developed.

Proof of Theorem 2.2

Recall that $a_n = \dfrac{\kappa_n}{\kappa_{n+1}}$ and so

$$\kappa_n = \left(\prod_{i=0}^{n} a_i \right)^{-1}. \tag{22}$$

By using Carleman inequality, see for instance [8], we get

$$\sum_{j=1}^{n} \left(\prod_{i=0}^{j} a_i \right)^{-1/j} \leq 2e \sum_{j=1}^{n} a_{j-1}^{-1}, \tag{23}$$

then by using (22) we obtain

$$\sum_{j=1}^{n} \left(\kappa_j \right)^{1/j} \leq 2e \sum_{j=1}^{n} a_j^{-1}. \tag{24}$$

By Lemma 4.1 we have $\kappa_n \geq (\mu_{2n})^{-1/2}$, and so (24) turns out to be

$$\sum_{j=1}^{n} \left(\mu_{2j}\right)^{-\frac{1}{2j}} \leq 2e \sum_{j=0}^{n-1} a_j^{-1}. \tag{25}$$

By using (20), inequality (25) becomes

$$\sum_{j=1}^{n} \left(\mu_{2j}\right)^{-\frac{1}{2j}} \leq 4e \left[\sum_{k=0}^{n} |e_k(x)|^2\right]^{1/2} \left[\sum_{k=0}^{n} |q_k(x)|^2\right]^{1/2}. \tag{26}$$

Since the probability measure is compactly supported, there exist $M > 0$ such that for every $n \in \mathbb{N}$

$$\mu_{2n} \leq M^{2n}. \tag{27}$$

From (27) we have for every $n \geq 1$

$$\left(\mu_{2n}\right)^{-\frac{1}{2n}} \geq \frac{1}{M},$$

then the series $\sum_{n=1}^{+\infty} \left(\mu_{2n+2}\right)^{-\frac{1}{2n}}$ is divergent and, because of (26), at least one of the series $\sum_{k=0}^{\infty} |e_k(x)|^2$ and $\sum_{k=0}^{\infty} |q_k(x)|^2$ is divergent. But since both sequences $\{e_k(x)\}$ and $\{q_k(x)\}$ are solutions of the same recursive equation (18), they both diverge, which finishes the proof of Theorem 2.2. □

References

[1] S. Barry, *The Classical Moment Problem as Self-adjoint finite difference Operator*, Advances in Mathematics **137** (1998), 82-203.
[2] R. Bellman, *Introduction to Matrix Analysis*, Society for Industrial and Applied Mathematics, (second edition), Philadelphia, 1997.
[3] G. Brian and D. Spencer, *The Classical Orthogonal Polynomials*, Word Scientific, 2016.
[4] H. Chebli, *Analyse Hilbertienne*, Centre de Publication Universitaire ,Tunis, 2001.
[5] R. Cools, F.Y. Kuo and D. Nuyens, *Constructing embedded lattice rules for multivariate approximation in* $[0,1]^s$, SIAM - Journal on Scientific Computing, 2009.
[6] R. Devore and G.G. Lorentz, *Constructive Approximation*, Springer, 1993.
[7] M. Foupouagnigni, W. Koepf, *Orthogonal Polynomials*, Tutorials, Schools, and Workshops in the Mathematical Sciences, 2nd AIMS-Volkswagen Stiftung Workshop, Douala, Cameroon, 2018.
[8] I. S. Gradshteyn, I. M. Ryzhik, *Table of Integrals, Series, and Products*, Alan Jeffrey, Editor and Daniel Zwillinger, (Seventh Edition), 2007.
[9] G.G. Lorentz, *Constructive Approximation: A Status Report*, Journal of Approximation Theory, 1986.
[10] Y. Nakatsukasa and L.N. Trefethen, *The Pyramid Algorithm for Approximating Polynomials and Rational Functions*, SIAM - Journal on Numerical Analysis, 2013.
[11] G. M. Phillips, *Interpolation and Approximation by Polynomials*, CMS Books in Mathematics, Springer, 2003.
[12] A. Pinkus, *Weierstrass and Approximation Theory*, Journal of Approximation Theory, **107**, (2000), 1–66.
[13] J. B. Prolla, *Weierstrass-Stone: The Theorem*, Peter Lang GmbH, Internationaler Verlag der Wissenschaften, Approximation and Optimization, 1993.
[14] A. Rezgui, *On Interpolation of* L^2 *function*. International Journal of Statistics and Probability, **14(2)**, (2017), 1–15.
[15] A. Rezgui, *Polynomial Interpolation of Normal Conditional Expectation*, The Australian Journal of Mathematical Analysis and Applications, **8(3)**, (2019), 75–82.
[16] T.J. Rivlin, *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory*, Bulletin of the American Mathematical Society, 1990.
[17] K. Schmüdgen, it The Moment Problem, Springer International Publishing AG, (2017), 85–89.
[18] M. H. Stone, *Applications of the theory of Boolean rings to general topology*, Transactions of the American Mathematical Society, 41, **3**, (1937), 375–481.
[19] L.N. Trefethen *Approximation Theory and Approximation Practice*, SIAM, 2013.
[20] H. Widom, H. S. Wilf, *Small eigenvalues of large Hankel matrices*, Proc. Am. Math. Soc. **17**, (1966), 338—344.
[21] Y. Xu and D. Zeng, *Approximation theory and methods for structured tensor data*, SIAM, 2017.
[22] M. Zhu, Y. Chen, N. Emmart, C. Weems, *The smallest eigenvalue of large Hankel matrices*, Appl. Math. Comput. **334**, (2018), 375—387.