

Pointwise Convergence of Jacobi Polynomials

Ekaterina Kruglov

School of Science

Thesis submitted for examination for the degree of Master of Science in Applied Mathematics.

Espoo 26.03.2018

Thesis supervisors:

Senior Lecturer Harri Hakula (Aalto University)
Prof. Michael Hanke (KTH)

Author: Ekaterina Kruglov

Title: Pointwise Convergence of Jacobi Polynomials

Date: 26.03.2018

Language: English

Number of pages: 3+46

Department of Mathematics

Professorship:

Supervisors: Senior Lecturer Harri Hakula (Aalto University)
Prof. Michael Hanke (KTH)

The goal of this thesis is to numerically study a pointwise Jacobi convergence theorem for piecewise analytic functions based on the theorem on Legendre error. The convergence rates were examined on the boundary, singularity, and interior points. Results revealed that pointwise error convergence rates depend on the point of singularity, Jacobi polynomial coefficients α and β , and, lastly, on type of the piecewise analytic function.

Keywords: Orthogonal polynomials, piecewise analytic functions, Jacobi polynomials, approximation theorem, Gibbs phenomenon

Contents

Abstract	ii
Contents	iii
1 Introduction	1
2 Theoretical Background	3
2.1 Definition of Orthogonality	3
2.2 Orthogonal Polynomials	4
2.2.1 Jacobi Polynomials	5
2.2.2 Legendre polynomials	7
2.2.3 Chebyshev polynomials of the first kind	8
2.2.4 Chebyshev polynomials of the second kind	10
2.2.5 Gegenbauer polynomials	11
2.3 Least Squares	12
2.4 Applications	15
2.4.1 Finite Element Method	15
2.4.2 Spectral Analysis	18
3 Previous Results	20
3.1 Gibbs Phenomenon	20
3.1.1 Fourier series	21
3.1.2 Resolution of Gibbs Phenomenon	22
3.2 Legendre polynomials convergence rates	24
3.2.1 Pointwise error estimate of the Legendre Expansion	24
4 Research question	26
4.1 Theorem on $e_f(p, x, \alpha, \beta)$ for $f(x) \in \Phi$, $\mu > -1$	27
5 Method	28
5.1 Calculation of Jacobi Expansion Coefficients	28
5.1.1 Parallelization of Numerical Integration	29
5.2 Pointwise Error Approximation	29
5.3 Approximation of Pointwise error convergence rates (Theorem 3 (1-4))	29
5.4 Approximation of the Constants (Theorem 3 (5-7))	30
6 Results	30
6.1 Example 1 : $\mu \in \{0, 2, 2.5\}$, $a = \frac{1}{2}$, $\alpha = \beta = -\frac{1}{2}$	30
6.1.1 Verification of Theorem 3 (1-4)	30
6.1.2 Verification of Theorem 3 (5-7)	35
6.2 Example 2 : $\mu = 0$, $a = \frac{1}{2}$, $\alpha = 1$, $\beta = 2$	40
6.3 Example 3 : $\mu = 0$, $a = \frac{1}{2}$, $\alpha = -\frac{1}{2}$, $\beta = \frac{1}{2}$	41
7 Discussion	43
References	44

1 Introduction

Study of polynomials is dating back thousands years ago. Existence of a quadratic equation and its analytical solution, for example, were established already by the ancient Greeks, though, in their solutions they have considered only positive roots [Borwein and Erdelyi, 1995]. Cubic equations were also known to the majority of ancient population such as Babylonians, Greeks, Chinese, Indians, and Egyptians [Guilbeau, 1930]. For centuries, mathematicians from around the world, including Hippocrates, Menaechmus and Archimedes, had been trying to solve cubic equation and to find its analytical solution, but without success. Although majority of the attempts were unsuccessful, some came very close to the correct solution. Only during the 16th century, the method of solving cubic equations was finally found [Borwein and Erdelyi, 1995].

The rapid progress in the field of science began approximately in the 17th century when many significant results were established in fields such as mathematics and physics. Moreover, as a result of scientific progress, scientists were facing new sets of problems. For instance, one of the important problems that mathematicians were encountered with and are facing even today, was to express complex functions in terms of simpler ones.

In the beginning of 18th century, Brooke Taylor formulated Taylor theorem and Taylor series, which, half a century later, were proclaimed as basic principles of differential calculus [Britannica]. The purpose of Taylor series is to represent an arbitrary function at point a as an infinite sum of the terms that are obtained from derivatives of this function at point a [Zwillinger, 2011, Courant and McShane, 2011]. General expression of Taylor series is given by

$$f(x) = f(a) + f'(a)(x - a) + \frac{1}{2!}f''(a)(x - a)^2 + \frac{1}{3!}(x - a)^3 + \dots \quad (1)$$

Of course, in order to obtain the derivative, one must know the function on a very small neighborhood around each point.

Almost a century after the introduction of Taylor series, Jean-Baptiste Joseph Fourier introduced the Fourier series [Britannica]. By infinite sum of sines and cosines, which is referred to as Fourier series, one can expand any periodic function $f(x)$ (see Section 3.1.1) [Courant and McShane, 2011]. Since many problems in science are based on periodic functions, such as heat conduction or any vibrating phenomena, Fourier series quickly became very useful.

The study of Taylor and Fourier series brought a significant step forward in the progress of mathematical analysis [Jackson, 1921]. Both Taylor and Fourier series require certain qualities of a function [Courant and McShane, 2011]. For example, Taylor series, by definition, is an infinite sum of the derivatives of a function at point a , thus it assumes that a function is infinitely differentiable. On the other hand, Fourier series requires a function to be integrable on the whole interval of the expansion. Therefore, the next step was to find an approximation of any continuous functions by simple functions.

In the middle of 19th century Weierstrass formulated a theorem stating that for any continuous function $f(x)$ on a closed interval $[a, b]$, one can find a polynomial $P(x)$ such that for any $\epsilon > 0$ we have $|f(x) - P(x)| < \epsilon$ [Jackson, 1921]. This theorem showed that any complex function could be represented in terms of polynomials which is a simple function that is completely specified by a finite set of coefficients. As a consequence of Weierstrass theorem, mathematicians begun to devote their interests into the study of orthogonal polynomials.

Orthogonal polynomials is a family of polynomials that share the property of orthogonality, i.e. any two different polynomials in a sequence are orthogonal with respect to inner product (for more detailed definition see Section 2.2). The most widely studied orthogonal polynomials, often referred to as classical orthogonal polynomials, are Hermite polynomials, Laguerre polynomials, Jacobi polynomials with its special cases Gegenbauer, Chebyshev, and Legendre polynomials [Szegő, 1939, Zwillinger, 2011, Zarowski, 2004].

Classical orthogonal polynomials arise from the study of differential equations. In particular, Sturm-Liouville type of equation

$$Q(x)\frac{d^2f}{dx^2} + L(x)\frac{df}{dx} + \lambda f = 0 \quad (2)$$

where $Q(x)$ and L is a quadratic and linear polynomials, respectively. Sturm-Liouville problem (*S-L problem*) is given by (2) with prescribed boundary conditions [Lakshminarayanan and Varadharajan, 2015]. The solution of *S-L problem* consists of f and the eigenvalue of boundary conditions, λ , and it includes singularities in the solution of f , unless we take specific values for λ . Thus, the solution is a series of polynomials p_0, p_1, \dots, p_n of some degree n , each corresponding to λ_n .

In addition to the approximation of solution for differential equations, such as (2), orthogonal polynomials are powerful tools for approximating difficult or expensive to compute functions and integral solutions, which makes them essential for numerical integration. For instance, Chebyshev and Legendre polynomials are applied in approximation theory, numerical integration, study of elliptic equations, and computational fluid mechanics [Zarowski, 2004].

Today orthogonal polynomials is an essential tool in many fields of mathematics and physics. Using classical orthogonal polynomials, one can quickly and easily approximate functions that are difficult and expensive to compute. Furthermore, orthogonal polynomials perform particularly well in approximating functions that are continuous on a given interval. Often, however, functions are discontinuous and this reduces the accuracy of polynomial approximation.

In this thesis, we sought to verify Jacobi convergence theorem that is based on the theorem of Legendre error by Babuška and Hakula [Babuška and Hakula, 2016]. The findings of our numerical tests stand in line with the hypothesis. In particular, each statement of the theorem was tested and verified several times for various cases which gives us the confidence to conclude that the hypothesis is correct. Our findings, thus, serve as a starting point for theoretical work in proving the Jacobi convergence theorem.

The rest of this thesis is organized as follows. In Section 2, we provide the definition of orthogonality, orthogonal polynomials, and least squares approximation. The section is concluded by two examples of applications. In Section 3, we discuss previous results of polynomial approximation such as Gibbs phenomenon. Additionally, we discuss Legendre polynomial expansion of a piecewise analytic function, results of which lead to the research question of this thesis. Section 4 summarizes the research question formulated as Jacobi expansion theorem. In Section 5, we describe the methods used to numerically verify the research question posted in Section 4. Section 6 is a demonstration of results for three different examples. The summary and discussion of the results can be found in Section 7.

2 Theoretical Background

2.1 Definition of Orthogonality

Majority of orthogonality and orthogonal polynomials definitions presented in this paper were collected from a famous book by Gabor Szego *Orthogonal Polynomials* published in 1939 [Szego, 1939]. Consider a vector space V . We define an inner product as a function (\cdot, \cdot) from $V \times V$ to \mathbb{C} [Borwein and Erdelyi, 1995]. In particular, given a non-decreasing and non-constant function $\alpha(x)$ on the closed interval $[a, b]$ that satisfies the following

$$\alpha(-\infty) = \lim \alpha(x) < \infty$$

and

$$\alpha(\infty) = \lim \alpha(x) < \infty$$

then the inner product, or scalar product, is given by

$$(f, g) := \int_a^b f(x)g(x)d\alpha(x) \quad (3)$$

The inner product (3) satisfies the following three properties [Borwein and Erdelyi, 1995]:

- $(f, f) > 0$ unless $f = 0$
- $(f, g) = \overline{(g, f)}$
- $(\alpha f + \beta g, h) = \alpha(f, h) + \beta(g, h)$

Using the definition of inner product, we can now define the norm of f

$$\|f\| = (f, f) = \int_a^b (f(x))^2 d\alpha(x) \quad (4)$$

For a fixed function $\alpha(x)$, we say that a real function $f(x)$ is orthogonal to a real function $g(x)$, where $x \in [a, b] \in \mathfrak{R}$, with respect to $d\alpha(x)$, if the inner product is equal to zero, i.e.

$$(f, g) = \int_a^b f(x)g(x)d\alpha(x) = 0 \quad (5)$$

Moreover, we say that functions $f(x)$ and $g(x)$ are orthonormal if

- $(f, g) = 0$ for any two real function $f(x)$ and $g(x)$
- $(f, f) = 1$ and $(g, g) = 1$

Furthermore, consider a nonnegative and Lebesgue measurable function $w(x)$ such that $\int_a^b dx > 0$. Function $w(x)$ is often referred to as a weight function on a given interval $[a, b]$. Then, if $\alpha(x)$ is absolutely continuous, we can write the distribution $d\alpha(x)$ in terms of the weight function i.e. $w(x)dx$ and, thus, obtain

$$(f, g) = \int_a^b f(x)g(x)w(x)dx \quad (6)$$

2.2 Orthogonal Polynomials

In this paper, we deal with one variable polynomials of the form

$$p(x) = c_0 + c_1x + c_2x^2 + \dots + c_mx^m \quad (7)$$

where c_i are arbitrary coefficients for all i . A polynomial of degree m is denoted by $p_m(x)$.

By orthogonalizing the sequence $(1, x, x^2, \dots)$ we obtain the classical orthogonal polynomials [Borwein and Erdelyi, 1995]. The classes of polynomials $p_n(x)$ on a closed interval $[a, b]$ that satisfy (8), are called orthogonal polynomials.

$$\int_a^b p_m(x)p_n(x)w(x)dx = \delta_{mn}c_n \quad (8)$$

In (8), we have a weight function $w(x)$ and the Kronecker delta δ_{mn} (9).

$$\delta_{mn} = \begin{cases} 1, & \text{if } m = n. \\ 0, & \text{if } m \neq n. \end{cases} \quad (9)$$

If $m \neq n$, the right hand side of (8) is equal to 0, hence c_n does not play any role. On the other hand, when $m = n$, we get $\delta_{mn} = 1$ and, therefore

$$c_m = \int_a^b [p_m(x)]^2 w(x) dx \quad (10)$$

which is equivalent to the norm of p_m . Thus, in case we have $c_m = 1$, the orthogonal polynomials are also orthonormal.

Orthogonal polynomials satisfy a three term recurrence relations [Zarowski, 2004]

$$p_{n+1}(x) = (A_nx + B_n)p_n(x) + C_np_{n-1}(x) \quad (11)$$

where

$$\begin{aligned} A_n &= \frac{p_{n+1,n+1}}{p_{n,n}} \\ B_n &= \frac{p_{n+1,n+1}}{p_{n,n}} \left(\frac{p_{n+1,n}}{p_{n+1,n+1}} - \frac{p_{n,n-1}}{p_{n,n}} \right) \\ C_n &= \frac{p_{n+1,n+1}p_{n-1,n-1}}{p_{n,n}^2} \end{aligned}$$

Weight function, $w(x)$, serves an important role in the definition of various polynomials. Its purpose is to give a weight to errors at different points on a given interval $[a, b]$ [Zarowski, 2004]. Hence, if $w(x) = c > 0$, where c is some constant, then errors across the whole interval receive the same importance. One example of polynomials with a constant weight function is Legendre polynomials, which are orthogonal with respect to $w(x) = 1$ (see Section 2.2.2). If one is interested in giving more weights to certain points on the interval, then the weight function should be chosen differently. For example, Chebyshev polynomials (Sections 2.2.3 and 2.2.4) are orthogonal with respect to a weight function that gives more weights to the boundary points of the interval $[-1, 1]$.

In order to better understand the role of a weight function, consider a step function (12)

$$f(x) = \begin{cases} -1, & \text{for } -1 \leq x < a, \\ 0, & \text{for } x = a, \\ 1, & \text{for } a < x \leq 1, \end{cases} \quad (12)$$

Figure 1a illustrates the approximation of (12) by Legendre and Chebyshev polynomials of 10th degree. While both polynomials perform equally well on the interior points, the boundary points are better approximated by the Chebyshev polynomials (see Figure 1b) [Zarowski, 2004].

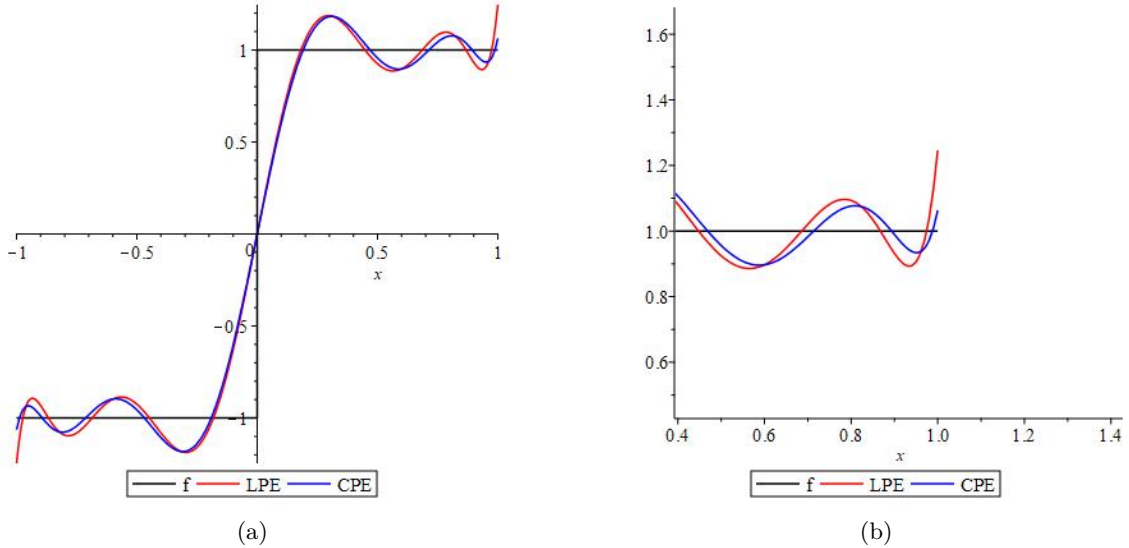


Figure 1: Comparison between Legendre and Chebyshev polynomials in approximating step function (a) and a close up view at a boundary point $x = 1$ (b)

2.2.1 Jacobi Polynomials

Jacobi polynomial $P_n^{(\alpha, \beta)}(x)$ is one of the classical orthogonal polynomials mentioned earlier and it is defined by

$$P_n^{(\alpha, \beta)}(x) = \frac{\Gamma(\alpha + n + 1)}{n! \Gamma(\alpha + \beta + n + 1)} \sum_{m=0}^n \binom{n}{m} \frac{\Gamma(\alpha + \beta + m + n + 1)}{n! \Gamma(\alpha + m + 1)} \left(\frac{x-1}{2}\right)^m \quad (13)$$

where $\Gamma(x)$ is called the Gamma function (14) and $\binom{n}{m}$ is a binomial coefficient ([Zwillinger, 2011] and [Szego, 1939]).

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt \quad (14)$$

An alternative definition of the Jacobi polynomials is given by the Rodrigues' formula

$$P_n^{(\alpha, \beta)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} (1-x)^\alpha (1+x)^\beta (1-x^2)^n \quad (15)$$

Jacobi polynomials are defined on the interval $[-1, 1]$ and are orthogonal with respect to the weight function

$$w(x) = (1-x)^\alpha (1+x)^\beta \quad (16)$$

where $\alpha, \beta > -1$, which ensures integrability [Szego, 1939]. Although Jacobi polynomials are not orthonormal, they can be normalized by

$$P_n^{(\alpha, \beta)}(1) = \binom{n + \alpha}{n} \quad (17)$$

Jacobi polynomials satisfy an important symmetry relation

$$P_n^{(\alpha,\beta)}(x) = (-1)^n P_n^{(\beta,\alpha)}(-x) \quad (18)$$

Like many orthogonal polynomials, Jacobi polynomials arise from solution of differential equation. In particular, $P_n^{(\alpha,\beta)}(x)$ is a solution to second order linear homogeneous differential equation, which is often referred to as Jacobi differential equation (19). Define $y = P_n^{(\alpha,\beta)}(x)$, then

$$(1-x^2) \frac{d^2 y}{dx^2} + [\beta - \alpha - (\alpha + \beta + 2)x] \frac{dy}{dx} + n(n + \alpha + \beta + 1)y = 0 \quad (19)$$

As discussed in Section 2.2, similarly to other orthogonal polynomials, Jacobi polynomials satisfy a three term recurrence relation (20) [Shen et al., 2011].

$$P_{n+1}^{\alpha,\beta}(x) = (a_n^{\alpha,\beta} x - b_n^{\alpha,\beta}) P_n^{\alpha,\beta}(x) - c_n^{\alpha,\beta} P_{n-1}^{\alpha,\beta}(x) \quad (20)$$

for $n \geq$ where

$$P_0^{\alpha,\beta}(x) = 0, \quad (21)$$

$$P_1^{\alpha,\beta}(x) = \frac{1}{2}(\alpha + \beta + 2)x + \frac{1}{2}(\alpha - \beta) \quad (22)$$

and

$$a_n^{(\alpha,\beta)} = \frac{(2n + \alpha + \beta + 1)(2n + \alpha + \beta + 2)}{2(n + 1)(n + \alpha + \beta + 1)} \quad (23)$$

$$b_n^{(\alpha,\beta)} = \frac{(2n + \alpha + \beta + 1)(\beta^2 - \alpha^2)}{2(n + 1)(n + \alpha + \beta + 1)(2n + \alpha + \beta)} \quad (24)$$

$$c_n^{(\alpha,\beta)} = \frac{(2n + \alpha + \beta + 2)(n + \alpha)(n + \beta)}{(n + 1)(n + \alpha + \beta + 1)(2n + \alpha + \beta)} \quad (25)$$

The recurrence relation allows us to evaluate Jacobi polynomials at any point $x \in [-1, 1]$ and it is also very useful in deriving other properties [Shen et al., 2011].

By alternating α and β , one can construct many different Jacobi polynomials. In Table 1 one can find a few examples of Jacobi polynomials with varying α and β for $n = 0, 1, 2$. Figure 2 illustrates the polynomials from Table 1 up to the fifth degree.

n	$\alpha = \beta = -\frac{1}{2}$	$\alpha = -\frac{1}{2}$ and $\beta = \frac{1}{2}$	$\alpha = \beta = \frac{1}{2}$
0	1	1	1
1	$\frac{1}{2}x$	$-\frac{1}{2} + x$	$\frac{3}{2}x$
2	$\frac{3}{4}(x-1)^2 + \frac{3}{2}x - \frac{9}{8}$	$\frac{3}{2}(x-1)^2 + \frac{9}{4}x - \frac{15}{8}$	$\frac{5}{2}(x-1)^2 + 5x - \frac{25}{8}$

Table 1: Examples of Jacobi polynomials $P_n^{(\alpha,\beta)}$ with different α and β for $n = 0, 1, 2$

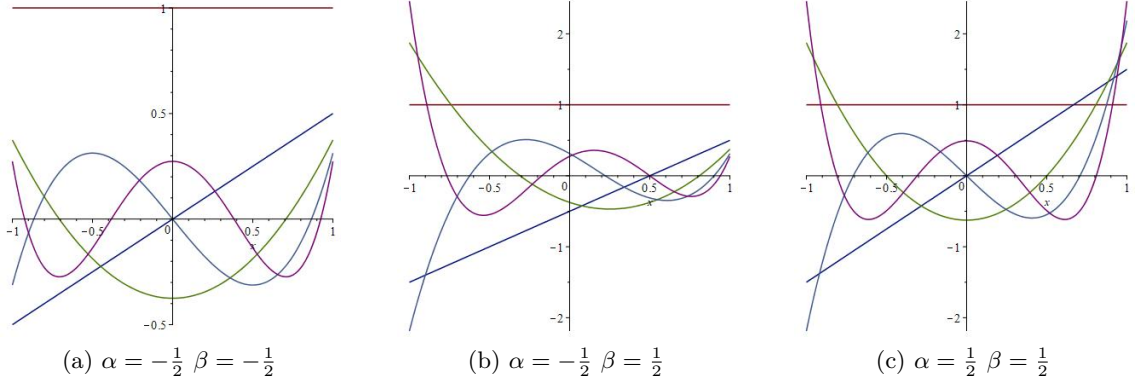


Figure 2: $P_n^{(\alpha, \beta)}(x)$ for $n = 0, \dots, 5$ and $x \in [-1, 1]$ with different α and β

Jacobi polynomials were introduced by and named after the German mathematician, Carl Gustav Jacob Jacobi (1804 - 1851) [Britannica]. During his career, Carl Jacobi made fundamental contributions to the study of elliptic functions which are now widely used in the field of mathematical physics. Additionally, Jacobi made several other significant contributions to multiple fields of mathematics including dynamics, differential equations, and number theory.

2.2.2 Legendre polynomials

Legendre polynomials are a special case of Jacobi polynomials where $\alpha = \beta = 0$ and they are often denoted by $P_n(x)$. Like Jacobi polynomials, $P_n(x)$ are orthogonal on $[-1, 1]$ with respect to the weight function (16) (i.e. $w(x) = 1$).

Legendre polynomials are solutions to the Legendre differential equations (26) where l is an integer [Zwillinger, 2011].

$$(1 - x^2) \frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + l(l + 1)y = 0 \quad (26)$$

The explicit expression of Legendre polynomials is given by (27)

$$P_n(x) = \frac{1}{2^n} \sum_{m=0}^{\frac{n}{2}} (-1)^m \binom{n}{m} \binom{2n - 2m}{n} x^{n-2m} \quad (27)$$

Moreover, we can express $P_n(x)$ using Rodrigues' formula

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n]$$

Legendre polynomials can be defined as coefficients for Taylor series expansion

$$\frac{1}{\sqrt{1 - 2xt + t^2}} = \sum_{n=0}^{\infty} P_n(x) t^n \quad (28)$$

from which we get the first two terms $P_0(x) = 1$ and $P_1(x) = x$. Further expansion of the Taylor series, gives us the *Bonnet's recursion formula*

$$(n + 1)P_{n+1}(x) = (2n + 1)xP_n(x) - nP_{n-1}(x) \quad (29)$$

As a result of (29) and the first two terms, $P_0(x)$ and $P_1(x)$, the Legendre polynomials can be generated recursively.

$$P_n(x) = 2^n \sum_{k=0}^n x^k \binom{n}{k} \binom{\frac{n+k-1}{2}}{n} \quad (30)$$

From (30) we get $P_n(x)$ for $n = 0, \dots, 4$

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

which are then illustrated in Figure 3.

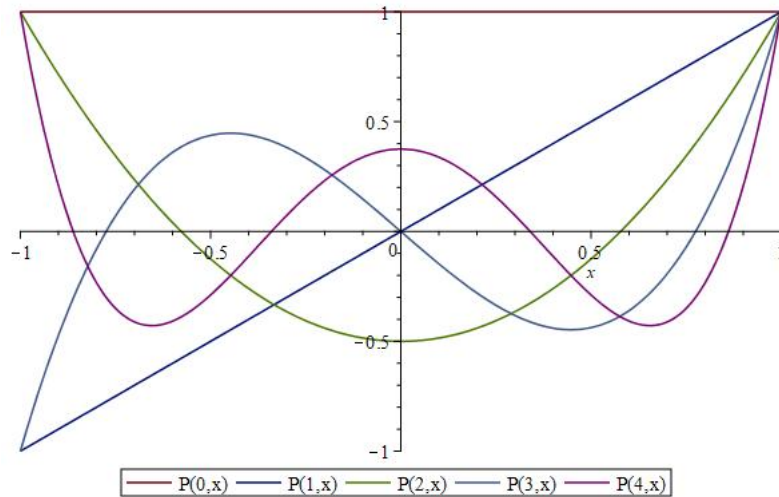


Figure 3: Legendre polynomials $P_n(x)$ for $n = 0, \dots, 4$

Legendre polynomials were first introduced by a French mathematician, Adrien-Marie Legendre in 18th century. During his career, Legendre made multiple significant contributions to various fields of mathematics such as study of elliptic functions, number theory, and the method of least squares. One of his most famous publications was the book *Elements de geometrie* published in 1794. This book was a leading elementary textbook for geometry in the following hundred years.

2.2.3 Chebyshev polynomials of the first kind

Chebyshev polynomials of the first kind are solutions to the Chebyshev differential equations of second order (31) where $|x| < 1$ and $a \in \mathfrak{R}$, constant [Zwillinger, 2011].

$$(1 - x^2) \frac{d^2 y}{dx^2} - x \frac{dy}{dx} + a^2 y = 0 \quad (31)$$

The general solution of the differential equation (31) can be written in a closed form

$$y = a_0 \cos(\alpha \sin^{-1} x) + \frac{a_1}{\alpha} \sin(\alpha \sin^{-1} x)$$

After performing a change of variables, we get an equivalent form of the solution

$$y = b_1 T_\alpha(x) + b_2 \sqrt{1-x^2} U_{\alpha-1}(x) \quad (32)$$

where $T_n(x)$ and $U_n(x)$ are Chebyshev polynomials of first and second kind, respectively (see Section 2.2.4 for the latter).

Chebyshev polynomials of the first kind are a special case of Jacobi polynomials with $\alpha = \beta = -\frac{1}{2}$ that are orthogonal on $[-1, 1]$ with respect to the weight function (33) ([Szegő, 1939] and [Zwillinger, 2011])

$$w(x) = (1-x^2)^{-\frac{1}{2}} \quad (33)$$

The explicit expression (34) and Rodrigues's formula (35) of Chebyshev polynomials of the first kind

$$T_n(x) = \frac{n}{2} \sum_{m=0}^{\frac{n}{2}} (-1)^m \frac{(n-m-1)!}{m!(n-2m)!} (2x)^{n-2m} \quad (34)$$

$$T_n(x) = \frac{\sqrt{\pi(1-x^2)}}{(-2)^n \Gamma(n+\frac{1}{2})} \frac{d^n}{dx^n} [(1-x^2)^{n-\frac{1}{2}}] \quad (35)$$

Just like Jacobi and Legendre polynomials, Chebyshev polynomials of the first kind have a recurrence relation that allows us to express T_{n+1} in terms of T_n and T_{n-1} , (36).

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (36)$$

for $n \geq 1$ where $T_0(x) = 1$ and $T_1(x) = x$.

It is also worth mentioning an interesting relation between T_n and trigonometric functions. Consider

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(\theta)\cos(n\theta)$$

If we let $\theta = \arccos(x)$, recall that $\cos(\arccos(x)) = x$ and move the second term on the left-hand side to the right, we get

$$\cos((n+1)\arccos(x)) = 2x\cos(n\arccos(x)) - \cos((n-1)\arccos(x))$$

which is equivalent to the recurrence relation (36). Thus, we can write

$$T_n(x) = \cos(n\arccos(x)), \quad \text{where } n \geq 0 \text{ and } x \in [-1, 1] \quad (37)$$

One of the direct consequences of (37) is that the Gauss-type quadrature nodes and weights can be derived explicitly. The explicit definition eliminates the potential loss of accuracy when performing numerical procedure for large numbers [Gottlieb and Orszag, 1977]. The first five Chebyshev polynomials of the first kind are:

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \end{aligned}$$

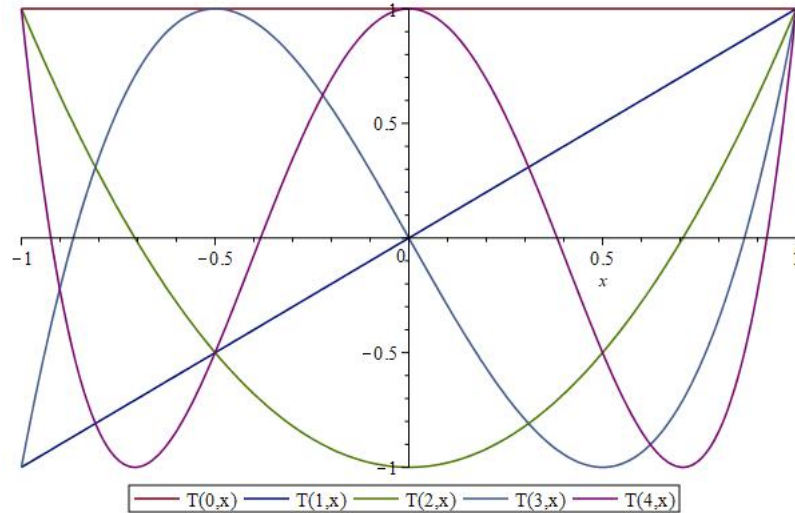


Figure 4: Chebyshev polynomials of the first kind, $T_n(x)$ for $n = 0, \dots, 4$

and their plot can be seen in the Figure 4.

Chebyshev polynomials were introduced and named after the Russian mathematician, Pafnuty Chebyshev (1821-1894). In addition to his contribution to the proof of prime number theorem, Chebyshev is famous for his work with approximation of functions. During his career, Chebyshev founded Saint Petersburg's mathematical school [Britannica]. Today, Chebyshev polynomials are extensively applied in the study of spectral methods.

2.2.4 Chebyshev polynomials of the second kind

Chebyshev polynomials of second kind serve as a part of general solution to the Chebyshev differential equation (31) and they are also a solution to (38).

$$(1 - x^2) \frac{d^2 y}{dx^2} - 3x \frac{dy}{dx} + n(n + 2)y = 0 \quad (38)$$

Chebyshev polynomials of second kind, $U_n(x)$, are also a special case of Jacobi polynomials with $\alpha = \beta = \frac{1}{2}$. $U_n(x)$ are orthogonal on $[-1, 1]$ with respect to the weight function $w(x) = (1 - x^2)^{\frac{1}{2}}$. The explicit expression (39) and Rodrigues's formula (40) of Chebyshev polynomials of the second kind

$$U_n(x) = \sum_{m=0}^{\frac{n}{2}} \frac{(-1)^m (n - m)!}{m! (n - 2m)!} (2x)^{n-2m} \quad (39)$$

$$U_n(x) = \frac{(-1)^n (n + 1) \sqrt{\pi}}{(1 - x^2)^{\frac{1}{2n+1}} \Gamma(n + \frac{3}{2})} \frac{d^n}{dx^n} [(1 - x^2)^{n+\frac{1}{2}}] \quad (40)$$

Additionally, Chebyshev polynomials of the second kind satisfy a recurrence relation

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x) \quad (41)$$

where $U_0(x) = 1$ and $U_1(x) = 2x$. Similarly to the first kind, $U_n(x)$ satisfy a special trigonometric relation for $x \in [-1, 1]$ and $n = 0, 1, 2, \dots$

$$U_n(x) = \frac{\sin[(n + 1)\cos^{-1}(x)]}{\sin[\cos^{-1}(x)]} \quad (42)$$

Below are listed $U_n(x)$ for $n = 0, \dots, 4$ and then illustrated in Figure ??

$$\begin{aligned} U_0(x) &= 1 \\ U_1(x) &= 2x \\ U_2(x) &= 4x^2 - 1 \\ U_3(x) &= 8x^3 - 4x \\ U_4(x) &= 16x^4 - 12x^2 + 1 \end{aligned}$$

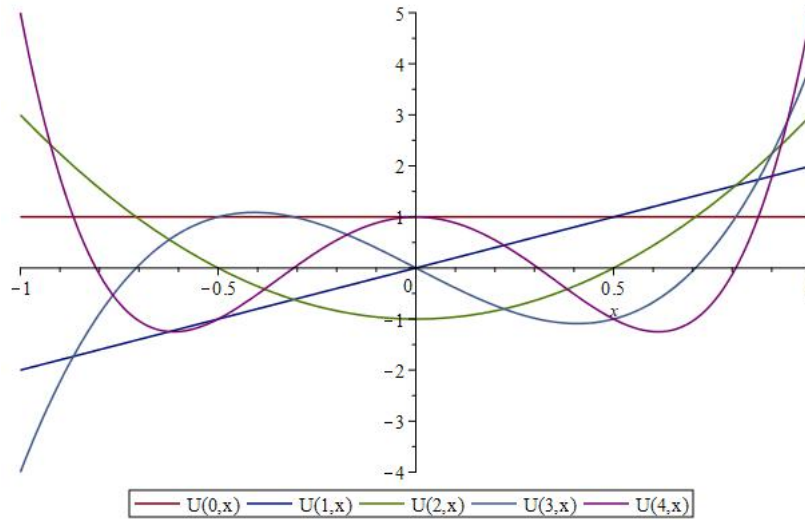


Figure 5: Chebyshev polynomials of the second kind, $U_n(x)$, for $n = 0, \dots, 4$

2.2.5 Gegenbauer polynomials

Gegenbauer polynomials, denoted by $C_n^\lambda(x)$, are also a special case of Jacobi polynomials where $\alpha = \beta = \lambda - \frac{1}{2}$ [Szegő, 1939]. $C_n^\lambda(x)$ are orthogonal on $[-1, 1]$ with respect to the weight function (43)

$$w(x) = (1 - x^2)^{\lambda - \frac{1}{2}} \quad (43)$$

In addition to being a special case of $P_n^{\alpha, \beta}(x)$, $C_n^\lambda(x)$ is generalization of $P_n(x)$ if we set $\lambda = \frac{1}{2}$, i.e. $C_n^{\frac{1}{2}}(x) = P_n(x)$. When $\lambda = 0$, $C_n^\lambda(x)$ vanish, which means that we cannot obtain $T_n(x)$ from Gegenbauer polynomials. On the other hand, $U_n(x)$ are easily obtained from $C_n^\lambda(x)$ by letting $\lambda = 1$ [Lakshminarayanan and Varadharajan, 2015]

Similarly to Jacobi polynomials and its special cases, $C_n^\lambda(x)$ are a solutions to differential equation. Let $y = C_n^\lambda(x)$, then

$$(1 - x^2) \frac{d^2 y}{dx^2} - (2\lambda + 1) \frac{dy}{dx} + n(n + 2\lambda)y = 0 \quad (44)$$

Below are explicit expression (45) and Rodrigues's formula (46) of Gegenbauer polynomials [Lakshminarayanan and Varadharajan, 2015]

$$C_n^\lambda(x) = (-1)^n \sum_{m=0}^{\frac{n}{2}} \binom{-\lambda}{n-m} \binom{n-m}{m} (2x)^{n-2m} \quad (45)$$

$$C_n^\lambda(x) = \frac{(-1)^n \Gamma(\lambda + \frac{1}{2}) \Gamma(n + 2\lambda)}{2^n n! \Gamma(2\lambda) \Gamma(n + \lambda + \frac{1}{2})} (1 - x^2)^{-\lambda + \frac{1}{2}} \frac{d^n}{dx^n} [(1 - x^2)^{n + \lambda - \frac{1}{2}}] \quad (46)$$

Additionally, Gegenbauer polynomials satisfy a three-term recurrence relation (47)

$$(n + 1)C_{n+1}^\lambda(x) = 2(\lambda + n)x C_n^\lambda(x) + (2\lambda + n - 1)C_{n-1}^\lambda(x) \quad (47)$$

Below are listed $C_n^\lambda(x)$ for $n = 0, \dots, 3$ in their general form.

$$\begin{aligned} C_0^\lambda(x) &= 1 \\ C_1^\lambda(x) &= 2\lambda x \\ C_2^\lambda(x) &= -\lambda + 2\lambda(1 + \lambda)x^2 \\ C_3^\lambda(x) &= -2\lambda(1 + \lambda)x + \frac{4}{3}\lambda(1 + \lambda)(2 + \lambda)x^3 \end{aligned}$$

The plots of Gegenbauer polynomials for $\lambda = \frac{1}{2}$ and $\lambda = 1$ can be found in Figure 3 and Figure 5, respectively. Figure 6 illustrates $C_n^\lambda(x)$ for $\lambda = \frac{3}{2}$ and 2.

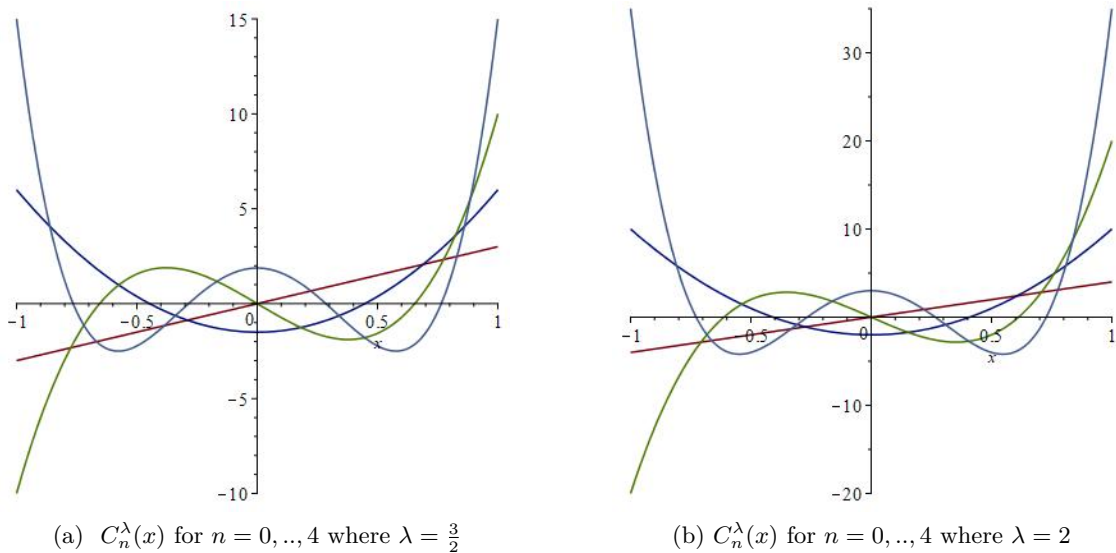


Figure 6: Gegenbauer polynomials, $C_n^\lambda(x)$, for $n = 0, \dots, 4$ where $\lambda = \frac{3}{2}$ (6a) and $\lambda = 2$ (6b)

Gegenbauer polynomials were introduced by the Austrian mathematician, Leopold Gegenbauer (1849 - 1903) [Lakshminarayanan and Varadharajan, 2015]. His interests involved number theory, complex analysis, and theory of integration.

2.3 Least Squares

Besides previously mentioned approximation methods, there exists a method that seeks to find approximation $p(x)$ of $f(x)$ by reducing the average error over the interval of approximation. This method is referred to as *Least Squares Approximation*. Least squares is equivalent to the expansion in orthogonal polynomials and hence it is used to define polynomial approximations [Scott, 2011].

Example: Consider $f(x) = e^x$ and $p(x) = \alpha_1 + \alpha_2 x$ on interval $[-1, 1]$. We want to find α_1 and α_2 that minimize

$$g(\alpha_0, \alpha_1) = \int_{-1}^1 [e^x - \alpha_0 - \alpha_1 x]^2 dx \quad (48)$$

i.e. we want to find α_1 and α_2 such that

$$\frac{\partial g}{\partial \alpha_0} = 0 \text{ and } \frac{\partial g}{\partial \alpha_1} = 0 \quad (49)$$

Figure 7 shows the approximation of function $f(x)$ using Least square method, Legendre polynomials, and Chebyshev polynomials. The smallest average error was achieved by the least squares method.

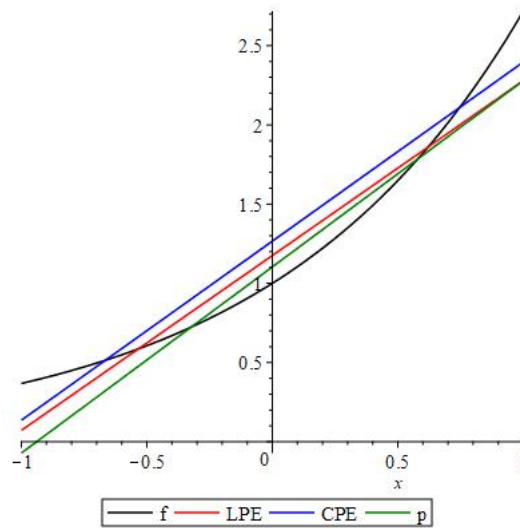


Figure 7: Approximation of $f(x) = e^x$ by Least squares (p), Legendre (LPE) and Chebyshev (CPE) polynomials.

In general, given a function $f(x)$ and $p(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_n x^n$, we want to minimize

$$g(\alpha_0, \alpha_1, \dots, \alpha_n) = \int_{-1}^1 [f(x) - p(x)]^2 dx$$

such that

$$\frac{\partial g}{\partial \alpha_i} = 0, \text{ for } i = 0, 1, \dots, n \quad (50)$$

As a result, we have a system of linear equation

$$A\alpha = b$$

where A is a $(n+1) \times (n+1)$ matrix and b is a vector with $n+1$ rows.

$$A_{i,j} = \frac{1}{i+j+1}$$

$$b_i = \int_{-1}^1 f(x)x^i dx$$

for $i, j = 0, 1, \dots, n + 1$.

Often, however, least square approximation using monomial polynomials results in ill-conditioned matrix A , such that increasing polynomial degree increases the rounding errors. One can replace the monomial polynomials with classical orthogonal polynomials. For example, one can set

$$p(x) = \sum_{i=0}^N \alpha_i P_i(x)$$

where $P_i(x)$ is Legendre polynomial and solve for α_i that minimizes $g(\alpha_0, \alpha_1, \dots, \alpha_N)$. By replacing monomial polynomials with orthogonal, we obtain a diagonal matrix A that is no longer ill-conditioned and the coefficient computations become faster and easier [Zarowski, 2004].

We start by constructing an orthonormal set of vectors $\{p^1, \dots, p^k\}$ from linearly independent set $\{v^1, \dots, v^k\}$ that spans subspace V_k of V using induction

$$p^i = t_i v^i + w^i$$

where $t_i \neq 0$ and $w^i \in V_{i-1}$. Then, given any vector $v_{k+1} \in V$ linearly independent from v_i for $i = 1, \dots, k$, we can construct the closest element of V_k to v^{k+1} by using least square [Scott, 2011]. Since $\{v^1, v^2, \dots, v^k\} \in V$ are linearly independent, we have $v^{k+1} \notin V_k$, and hence we define

$$p^{k+1} = \frac{1}{\|v^{k+1} - L_k^S v^{k+1}\|_2} (v^{k+1} - L_k^S v^{k+1})$$

where $L_k^S f$ is the *least square projection* defined by

$$L_k^S f = \sum_{i=1}^k (f, p^i) p^i \quad (51)$$

for $f \in V$.

Recall the definition of orthogonal polynomials (8) where $c_n = 1$. The polynomials are linearly independent and each polynomial $p_i(x)$ of degree i can be written in the form

$$p_i(x) = a_i x^i + q_i(x) \quad (52)$$

where $a_i \neq 0$ and the degree of $q_i(x)$ is $i - 1$. As a result, we formed a space P_n of degree n from the set of polynomials $p_i(x)$ for $i = 1, \dots, n$.

Following (52), we can now construct orthogonal polynomials. Taking $v^k = v^k(x) = x^{k-1}$ we obtain the least-squares projection that is defined for any $f \in V$ by (51). The next theorem is very useful in defining orthogonal polynomials in terms of least squares projection [Scott, 2011].

Theorem 1. *Given any $f \in V$,*

$$(f - L_n^S f, q) = 0$$

for all polynomials q of degree n , and

$$\|f - L_n^S f\|_2 = \min \|f - q\| \quad (53)$$

Theorem 1 leads to an important result which is useful in defining orthogonal polynomials.

Result: We have

$$\|f - L_n^S f\|_2 = 0 \quad (54)$$

if and only if $f \in P_n$.

By definition (52) of a polynomial of degree i , we note that x^{i+1} is not a polynomial and $x^{n+1} \notin P_n$. Thus, the orthogonal polynomials can be defined by

Definition 1.

$$P_{n+1} = a_{n+1}(x^{n+1} - L_n^S x^{n+1}) \quad (55)$$

where the coefficient

$$a_{n+1} = \frac{1}{\|x^{n+1} - L_n^S x^{n+1}\|_2} \quad (56)$$

is well-defined.

Definition 1 ensures that polynomials satisfy the orthogonality conditions (57)

$$\begin{cases} (P_{n+1}, P_{n+1}) & = 1 \\ (P_{n+1}, P_j) & = 0 \end{cases} \quad (57)$$

2.4 Applications

2.4.1 Finite Element Method

The finite element method (FEM) is one of the most widely used numerical method to solve various engineering and mathematical physics problems such as structural analysis, heat transfer, and fluid flow. These problems are often expressed in terms of partial differential equations (PDE) with defined boundary conditions. Frequently such PDEs cannot be solved analytically and therefore, their solutions are approximated by using different types of discretization methods.

Consider a function u that represents the dependent variable in PDE such as temperature, pressure, electric potential [Szabo and Babuška, 2011]. In order to approximate u , we use a trial function u_n (58) which is defined by linear combination of basis functions $\phi_i(x)$ (59)

$$u_n = \sum_{i=1}^n a_i \phi_i(x) \quad (58)$$

$$\phi_i(x) = x_i(l - x) \quad (59)$$

where l is the length of a rod, beam, or any other matter, and where $\phi_i(0) = \phi_i(l) = 0$. Finally, the goal of FEM is to minimize

$$I = \frac{1}{2} \int_0^l (k(u' - u_n')^2 + c(u - u_n)^2) dx \quad (60)$$

The family of functions that can be expressed by (58) is denoted by $S(I)$ where n represents the dimension of S . In order to construct S , we partition the solution into finite number of elements and on them we then define polynomial basis functions. The piecewise polynomials work well for characterization of basis function because of the following reasons:

- Piecewise polynomials are nonzero over a few elements only
- Advantageous in implementation
- Good approximation properties
- Linearly independent

There are two major versions of FEM: h-version and p-version [Babuška et al., 1981, Szabo and Babuška, 2011]. By using the first version, we approximate the solution by controlling for maximum diameter h of finite elements. In other words, with h-version, we obtain a more accurate solution by decreasing h and, thus, increasing the number of elements. In p-version, we do not change the size or number of the elements but instead, we increase the degree p of the polynomials. By increasing the polynomial order, we increase the complexity of shape functions. There also exists h-p version, which is a combination of the h- and p-versions. With h-p version, one increases number and degree of elements simultaneously, which can be done either uniformly or selectively [Babuška and Guo, 2006].

In the p-version of FEM, Lagrange and hierarchic shape functions (scaled integrals of Legendre polynomials) are frequently used in defining the basis functions.

- **Lagrange Shape functions:** As described earlier, FEM approximates the solution by discretization of the problem into a finite number of elements that are described by Lagrange Shape functions. The elements can be of a triangular or rectangular shape.

Consider an arbitrary triangle e with vertex j and coordinates $\{(x_i, y_i)\}$ where $i = 1, 2, 3$. Define a domain occupied by a triangle e by Ω_e . The linear shape function satisfies

$$N_j(x_k, y_k) = \delta_{j,k} \quad (61)$$

where the form of the 1st-order polynomial is

$$N_j(x, y) = a + bx + cy$$

where $(x, y) \in \Omega_e$. Combining the form function with the conditions (61), we get the following system of linear equations.

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & x_j & y_j \\ 1 & x_k & y_k \\ 1 & x_l & y_l \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

where $k \neq j \neq l$.

In order to construct higher-order Lagrangian shape functions, we use p -degree polynomials.

$$N_j(x, y) = \sum_{i=1}^{n_p} a_i q_i(x, y) = \mathbf{a}^T \mathbf{q}(\mathbf{x}, \mathbf{y})$$

where

$$\mathbf{q}^T(\mathbf{x}, \mathbf{y}) = [1, x, y, x^2, xy, y^2, \dots, y^p]$$

and n_p which is a number of monomial terms defined by

$$n_p = \frac{1}{2}(p+1)(p+2)$$

Recall, in the p-version of FEM, the number of elements does not increase and remains the same. Thus, we achieve a more accurate solution by increasing the order of shape functions (see Figure 8)

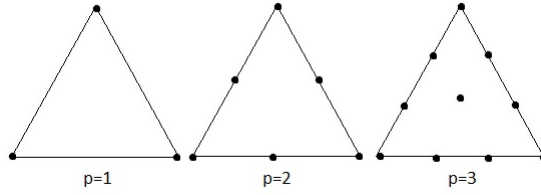


Figure 8: Lagrangian Shape Function on Triangle for $p=1,2,3$

- **Hierarchic shape functions:** In hierarchic shape functions, the basis of degree $p + 1$ is obtained as a correction to that of degree p , which explains the origins of its title. The hierarchic shape functions are constructed on Legendre polynomials which are less susceptible to round-off error accumulation at higher order than Lagrangian basis. Similarly to the Lagrangian, hierarchic shape-functions can be constructed on triangles and rectangles, here we cover the construction of latter.

The hierarchical polynomial of order p consists of four vertex shape functions ((62),(63), edge functions (64), and interior shape functions ((70), (71)).

$$N_{i,j}^1 = \hat{N}_i(\xi)\hat{N}_i(\nu) \quad (62)$$

where $i, j = 1, 2$ and

$$\hat{N}_1(\xi) = \frac{1 - \xi}{2}, \quad \hat{N}_2(\xi) = \frac{1 + \xi}{2} \quad (63)$$

There are $4(p-1)$ shape functions associated with the midside nodes $(3, 1), (2, 3), (3, 2)$, and $(1, 3)$:

$$N_{3,1}^k(\xi, \nu) = \hat{N}_1(\nu)\hat{N}^k(\xi) \quad (64)$$

$$N_{2,3}^k(\xi, \nu) = \hat{N}_2(\nu)\hat{N}^k(\xi) \quad (65)$$

$$N_{1,3}^k(\xi, \nu) = \hat{N}_1(\xi)\hat{N}^k(\nu) \quad (66)$$

$$N_{3,1}^k(\xi, \nu) = \hat{N}_2(\xi)\hat{N}^k(\nu) \quad (67)$$

$$(68)$$

where $k = 1, 2, \dots, p$ and $\hat{N}^k(\xi)$ are one-dimensional hierarchical shape functions defined by

$$\hat{N}^k(\xi) = \sqrt{\frac{2k-1}{2}} \int_{-1}^{\xi} P_{k-1}(\sigma) d\sigma \quad (69)$$

where $P_{k-1}(\sigma)$ is the Legendre polynomial of degree $k-1$. The linear function (63) ensure continuity of the basis by "blending" edge functions (69).

There are $\frac{1}{2}(p-2)(p-3)$ interior functions for $p \geq 4$ with the centroid, Node(3,3). The first internal function is

$$N_{3,3}^{4,0,0} = (1 - \xi^2)(1 - \nu)^2 \quad (70)$$

and the remaining shape functions are defined by product of the first internal function and the Legendre polynomials.

$$N_{3,3}^{5,1,0} = N_{3,3}^{4,0,0} P_1(\xi), \quad (71)$$

$$N_{3,3}^{5,0,1} = N_{3,3}^{4,0,0} P_1(\nu), \quad (72)$$

$$N_{3,3}^{6,2,0} = N_{3,3}^{4,0,0} P_2(\xi), \quad (73)$$

$$N_{3,3}^{6,1,1} = N_{3,3}^{4,0,0} P_1(\nu) P_1(\xi), \quad (74)$$

$$N_{3,3}^{6,0,2} = N_{3,3}^{4,0,0} P_2(\nu), \dots \quad (75)$$

$$(76)$$

In order to maintain continuity, these functions vanish on the element boundary, due to the special properties of Legendre polynomials.

Finally, the interpolant $U(\xi, \nu)$ is constructed by the linear combination on vertex, edge, and interior shape functions as follows

$$U(\xi, \nu) = \sum_{i=1}^2 \sum_{j=1}^2 c_{i,j}^1 N_{i,j}^1 + \sum_{k=1}^p \left[\sum_{j=1}^2 c_{3,j}^k N_{3,j}^k + \sum_{i=1}^2 c_{3,i}^k N_{i,3}^k \right] + \sum_{k=4}^p \sum_{\lambda+\mu=k-4}^2 c_{3,3}^{k,\lambda,\mu} N_{3,3}^{k,\lambda,\mu} \quad (77)$$

In total, there are $4 + 4(p-1) + \frac{1}{2}(p-2)(p-3)$ unknowns and shape functions in an approximation of degree p .

2.4.2 Spectral Analysis

Spectral Analysis is a class of techniques that is widely used in applied mathematics to numerically solve eigenvalue problems and ordinary and partial differential equations. Just like in FEM, the problem is written in terms of basis functions. In spectral method, the basis functions are usually nonzero over the whole domain which makes it a global method while FEM is local. In global approach, the value of a derivative at a certain point depends on the solution at all points in the domain and not only on the neighboring points. As a result, the spectral method usually gives the exact derivative of a function [Shen et al., 2011].

The choice of the trial function depends on the behavior of a problem. When dealing with a periodic problem, the trial functions are usually constructed by the Fourier series. Fourier series are very good in approximating continuous functions but, their accuracy and rate of convergence decreases significantly when dealing with discontinuities. When dealing with non-periodic problems, orthogonal polynomials such as Chebyshev and Legendre are preferred [Shen et al., 2011, Gottlieb and Orszag, 1977].

Consider a mixed initial boundary problem

$$\frac{\partial u(x, t)}{\partial t} = L(x, t)u(x, t) + f(x, t) \quad (78a)$$

$$B(x)u(x, t) = 0 \quad (78b)$$

$$u(x, 0) = g(x) \quad (78c)$$

where

$$\begin{aligned}
D & - \text{spatial domain} \\
\partial D & - \text{boundary} \\
L(x, t) & - \text{Linear differential operator} \\
B(x) & - \text{Linear (time-independent) boundary operator}
\end{aligned}$$

We assume that $u(x, t)$ is an element of a Hilbert space H with inner product (\cdot, \cdot) and norm $\|\cdot\|$. The solution $u(t) \in B$ where $B \in H$, satisfies the boundary condition (78b). Finally, L is an unbounded differential operator in a dense domain.

Problem (78a) has the following semi-discrete approximations

$$\frac{\partial u_N(x, t)}{\partial t} = L_N(x, t)u_N(x, t) + f_N(x, t) \quad (79)$$

where t and $u_N(x, t)$ belong to an N -dimensional subspace of B , B_N . L_N and f_N are given by

$$L_N = P_N L P_N \quad (80)$$

$$f_N = P_N f \quad (81)$$

where P_N is a projection operator from H onto B_N . As a result, there are two essential steps in formulation of spectral method:

- Choosing the approximation space B_N
- Choosing the projection operator P_N

Problems such as (78a) concern only one dependent variable u and one spatial variable x . In this paper, we will discuss only the Galerkin approximation method. Besides Galerkin, there are also Tau and Collocation approximation methods [Gottlieb and Orszag, 1977]. Considering problem (78a), the only difference between the three methods is in the treatment of boundary conditions. However, the difference becomes more pronounced when dealing with more complex problems.

A Galerkin approximation u_N of (78a) takes a form of truncated series

$$u_N(x, t) = \sum_{n=1}^N a_n(t)\phi_n(x) \quad (82)$$

Similarly to the FEM, $\phi_n(x)$ are time independent functions that are also linearly independent, which ensures that $u_N(x, t)$ satisfies boundary conditions. The expansion coefficients $a_n(t)$ are determined by

$$\frac{d}{dt}(\phi_n, u_N) = (\phi_n, Lu_N) + (\phi_n, f) \quad (83)$$

The projection operator is defined

$$P_N u(x) = \sum_{n=1}^N \sum_{m=1}^N p_{nm}(\phi_m, u)\phi_n(x) \quad (84)$$

where p_{nm} are the elements of inverse of $N \times N$ matrix with (ϕ_n, ϕ_m) elements.

As mentioned previously, the two most common choices of basis functions in spectral method are Fourier series and Chebyshev polynomials. First, consider Fourier series, then our domain is $H = L_2(0, \pi)$ and basis function $\phi_n(x) = \sin(nx)$. Every function $u \in L_2(0, \pi)$ has a Fourier sine series that converges in L_2 norm, such that $\|u - P_N u\|_2 \rightarrow 0$ as $N \rightarrow \infty$. This, however, does not ensure the convergence of Galerkin approximation u_N to u .

For Chebyshev series, H is a space of functions on interval $[-1, 1]$. The choice of basis function depends on the type of a problem. If we seek to solve a general wave equation

$$u_t + u_x = f(x, t) \tag{85}$$

$$u(-1, t) = 0 \tag{86}$$

$$u(x, 0) = g(x) \tag{87}$$

then basis function are

$$\phi_n(x) = T_n(x) - (-1)^n T_0(x)$$

where $T_n(x)$ is the Chebyshev polynomial of degree n (Section 2.2.3).

3 Previous Results

3.1 Gibbs Phenomenon

As mentioned before, polynomials and Fourier series worked well in approximating continuous functions. Unfortunately, not all functions are continuous and therefore, the next question was whether polynomials and Fourier series could be used in approximating discontinuous functions.

Consider a square wave function

$$f(x) = \begin{cases} 1 & \text{if } x \in [i, i+1] \text{ and } i \text{ is even} \\ -1 & \text{if } x \in [i, i+1] \text{ and } i \text{ is odd} \end{cases} \tag{88}$$

Observe, function (88) is continuous on intervals but is discontinuous on the overlap i.e. when $x = 0, 1, 2, \dots$. When we attempt to approximate (88) by Fourier series, we obtain significantly larger errors at the points of discontinuities compared to the interior. Moreover, observe that the errors at discontinuous do not decrease as we increase the order of Fourier series (see Figure 9). The observed behavior is called *Gibbs phenomenon* which describes the overshoot of Fourier series approximation at the points of discontinuity and, the fact, that the overshoot does not decay with the increase of Fourier series terms [Oddy, 2015].

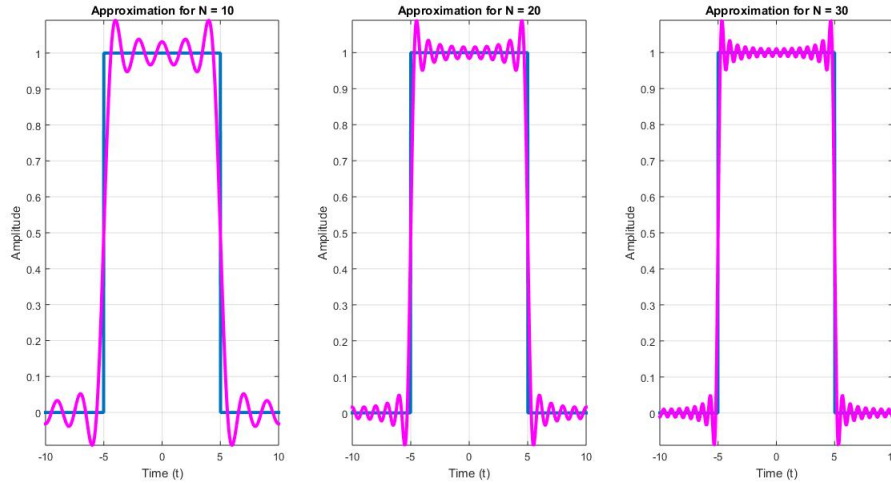


Figure 9: Fourier Series Approximation of the Wave Function (88)

The Gibbs phenomenon was first observed in the end of the 19th century by Michelson and Stratton who build a mechanic device that stored Fourier coefficients, known as harmonic analyzer. They published a paper where they presented the device and reconstructed functions from Fourier coefficients. One of these functions was function (88) which displayed what would later be called the Gibbs phenomenon.

Gibbs phenomenon, as one might correctly guess, brings unwanted consequences. For example, it causes artifacts in signal processing, MRI and image processing. Hence, the resolution of this phenomenon is highly motivated.

3.1.1 Fourier series

Consider a bounded periodic function $f(x)$ with period $2L$ that satisfies the following two conditions:

- $f(x)$ is continuous on the whole period or it has a finite number of jump discontinuities.
- $f(x)$ has a finite number of maxima and minima on any period.

Then, we can express $f(x)$ in terms of an infinite sum of sines and cosines (89), referred to as Fourier series [Zwillinger, 2011].

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right) \quad (89)$$

where a_n and b_n are called Fourier coefficients and are defined

$$a_n = \frac{1}{L} \int_{-L}^L f(x) \cos\left(\frac{n\pi x}{L}\right) dx$$

$$b_n = \frac{1}{L} \int_{-L}^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx$$

Fourier series are known to converge exponentially fast to any periodic function $f(x)$ that is analytic. Unfortunately, approximating a function that is either non periodic or

discontinuous yields in very poor results. Figure 9 illustrates Fourier series approximation of a periodic function that is discontinuous. The Fourier series approximation of a non periodic but continuous function $f(x) = x$ on the interval $-\pi \leq x \leq \pi$ is demonstrated in Figure 10

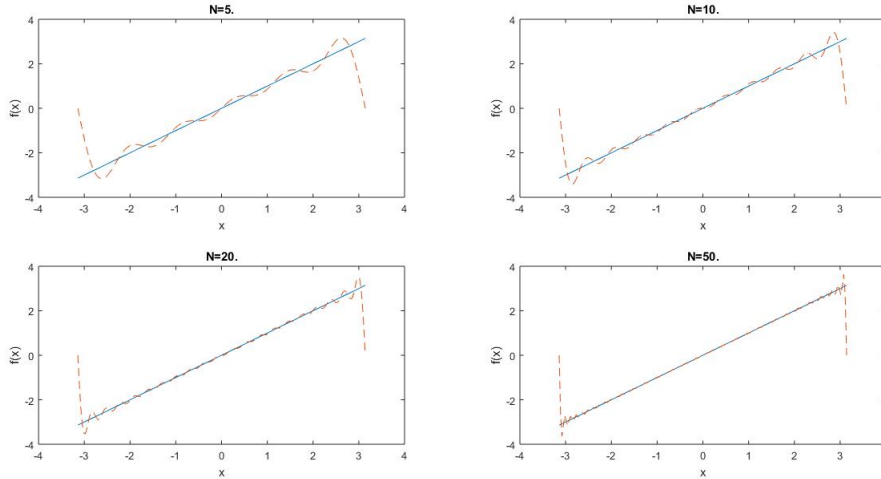


Figure 10: Fourier series approximation of continuous non-periodic function $f(x) = x$ on $[-\pi, \pi]$

Observe, the Fourier series provide a good approximation on the continuous part of the function, far away from the discontinuities. Overshoot occurs on the boundaries of the interval (in non-periodic function) or at the points of discontinuities/corners. The overshoot, however, is not random and does not grow in magnitude as we increase or decrease N . In fact, the overshoot is approximately 9% of the length of the jump discontinuity [Oddy, 2015].

3.1.2 Resolution of Gibbs Phenomenon

Resolution of the Gibbs phenomenon came almost two centuries after its discovery. David Gottlieb and Chi-Wang Shu successfully resolved the Gibbs phenomenon and almost completely removed the overshoot effect [Gottlieb and Shu, 1997]. Resolution of Gibbs phenomenon requires a different approximation of the Fourier coefficients. Instead of Fourier coefficients, Gottlieb and Shu, used a special case of Jacobi polynomials, the Gegenbauer polynomials (see Section 2.2.5).

First, we examine the non periodic analytic function on the interval $[-1, 1]$. Assuming the first $2N + 1$ Fourier coefficients $f_n(x)$, we, then, calculate the first $m + 1$ Gegenbauer coefficients, $\hat{g}_m^\lambda(x)$, (90)

$$\hat{g}_k^\lambda(x) = \frac{1}{h_k^\lambda} \int_{-1}^1 (1-x^2)^{\lambda-\frac{1}{2}} f_N(x) C_k^\lambda(x) dx \quad (90)$$

for $0 \leq k \leq m$, where h_k^λ is the normalization constant (91).

$$h_n^\lambda = \int_b^a w(x) [C_n^\lambda(x)]^2 dx = \pi^{\frac{1}{2}} C_n^\lambda(1) \frac{\Gamma(\lambda + \frac{1}{2})}{\Gamma(\lambda)(n + \lambda)} \quad (91)$$

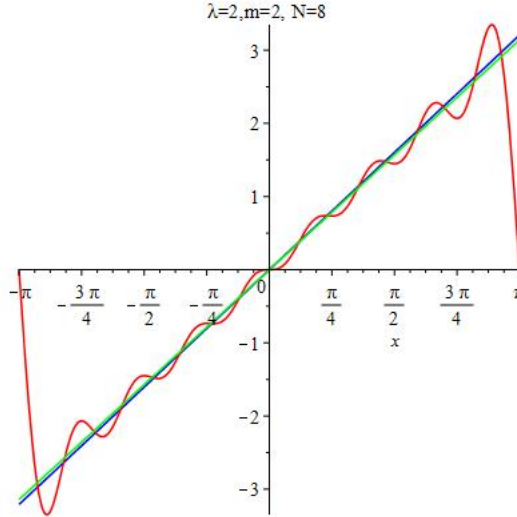


Figure 11: Approximation of a continuous non-periodic function $f(x) = x$ (Blue) by Fourier series (Red) and Fourier series with Gegenbauer coefficients (Green)

Using the constructed Gegenbauer coefficients in (90), we then calculate the Gegenbauer series approximation of $f(x)$

$$f(x) \approx \sum_{k=0}^m \hat{g}_k^\lambda C_k^\lambda(x) \quad (92)$$

Applying the Gegenbauer series approximation on the linear function $f(x) = x$ and setting $\lambda = 2$, $m = 2$, and $N = 8$, we note a significant improvement to the Fourier series approximation (see Figure 11)

We are now left to resolve the Gibbs phenomenon in the approximation of any piecewise analytic function $f(x)$ on $[-1, 1]$. Here we look at intervals $[a, b] \in [-1, 1]$ that do not include discontinuities. Each interval is approximated separately and the final results is the combination of interval approximations. For each interval $[a, b]$, we start by defining a local variable α such that $f(x)$ is analytic on $-1 \leq \alpha \leq 1$. Then, given

$$\epsilon = \frac{b-a}{2} \text{ and } \delta = \frac{b+a}{2}$$

we define

$$\alpha = \alpha(x) = \frac{x - \delta}{\epsilon}$$

Next, we calculate the first $m + 1$ Gegenbauer coefficients $\hat{g}_\epsilon^\lambda(l)$ of the Fourier series $f_N(\epsilon\alpha + \delta)$.

$$\hat{g}_\epsilon^\lambda(l) = \frac{1}{h_l^\lambda k} \int_{-1}^1 (1 - \alpha^2)^{\lambda - \frac{1}{2}} f_N(\epsilon\alpha + \delta) C_l^\lambda(x) d\alpha \quad (93)$$

Finally, using the above calculated coefficients, we can now construct Gegenbauer series to approximate

$$f(\epsilon\alpha + \delta) \approx \sum_{k=0}^m \hat{g}_\epsilon^\lambda(l) C_l^\lambda(\alpha) \quad (94)$$

It has been further shown that expanding the N^{th} Fourier coefficient into Gegenbauer polynomials produces very small errors. As a result, Gegenbauer procedure converges exponentially to any $f(x)$.

3.2 Legendre polynomials convergence rates

One of the most investigated and applied orthogonal polynomials are the Legendre polynomials which are frequently used in the interpolation and approximation theory, numerical integration, and special function theory. Legendre polynomials not only able to approximate globally smooth functions with a sufficiently small error but they are also very useful in approximating functions that are difficult to compute [Wang and Xiang, 2012]. Definition and special properties of Legendre polynomials one can find in Section 2.2.2. Recall, Gegenbauer polynomials are generalization of Legendre polynomials, which means that Legendre polynomials can also be used in resolving the Gibbs phenomenon.

Despite the wide range of applications, Legendre polynomials fall behind in the convergence rates. For example, when the number of coefficients is greater than 1 in (90) and (93), Legendre polynomials provide almost no improvement to the Fourier series approximation of non periodic or discontinuous functions.

Furthermore, the convergence rate of Legendre polynomials falls behind the convergence rate of Chebyshev polynomials. When comparing the decay rates of Legendre and Chebyshev coefficients, the decay rates of Chebyshev polynomials are always faster in approximation of discontinuous or non analytic functions (Figure 1). Even when the function is analytic on the interval and the coefficients decay exponentially, the decay of Legendre coefficients is still slightly slower [Wang and Xiang, 2012].

3.2.1 Pointwise error estimate of the Legendre Expansion

For years, mathematicians were interested in approximating functions that are difficult to compute. Many studies of function approximation by polynomials focused on measuring the global error convergence rates. On the other hand, pointwise error convergence has received very little attention. Therefore, the next step in the study of Legendre polynomials expansion was to examine the pointwise error estimation. Babuška and Hakula [Babuška and Hakula, 2016] addressed the class Φ of piecewise analytic functions with focus on $\mu = 0$

$$f(x) = \begin{cases} 0, & \text{for } -1 \leq x < a, \\ (x - a)^\mu, & \text{for } a \leq x \leq 1, \mu \geq -1, \end{cases} \quad (95)$$

their Legendre expansion

$$L_p(f)(x) = \sum_{k=1}^p a_k P_k(x) \quad (96)$$

and the error

$$e_f(p, x) = |f(x) - L_p(f)(x)| \quad (97)$$

as functions of a and μ based on the ideas of Wahlbin [Wahlbin, 1985]. The class of function Φ are frequently applied in structural mechanics where the boundary conditions are piecewise analytic functions [Babuška and Guo, 2006].

Consider a function (12) where a point of singularity is $a = \frac{1}{2}$. The pointwise error convergence rates α_1 and α_2 , were examined at two different points $x_1 = -1$ and $x_2 = -\frac{36}{75}$, respectively. Observe, x_1 is the boundary point of the interval $[-1, 1]$, while point x_2 is very close to the point of singularity, a . The results showed that the pointwise convergence rate at the boundary was slower than at the point of singularity, in particular, $\alpha_1 \approx -\frac{1}{2} \geq \alpha_2 \approx -1$. Figure 12 illustrates the pointwise convergence rates and the respective α at points x_1 and x_2 .

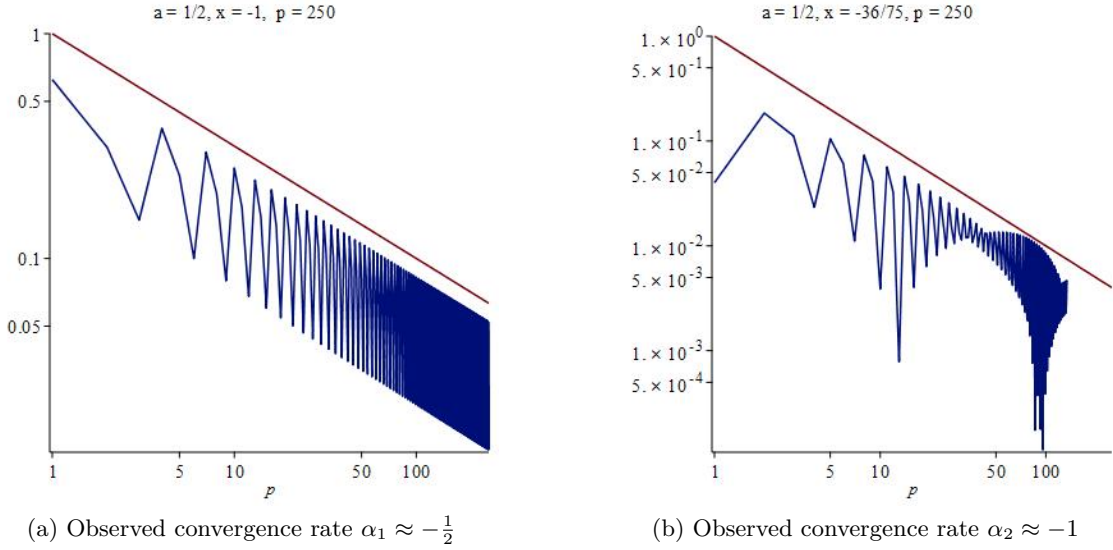


Figure 12: Pointwise error convergence rates of the Legendre polynomial expansion of function (12) with $a = \frac{1}{2}$ at points -1 (12a) and a (12b)

Next, consider points x_1 and x_2 and a new point of singularity $a = \frac{1}{7}$. Similar to the previous result, the pointwise error convergence rate is greater at points far from the boundary (see Figure 13).

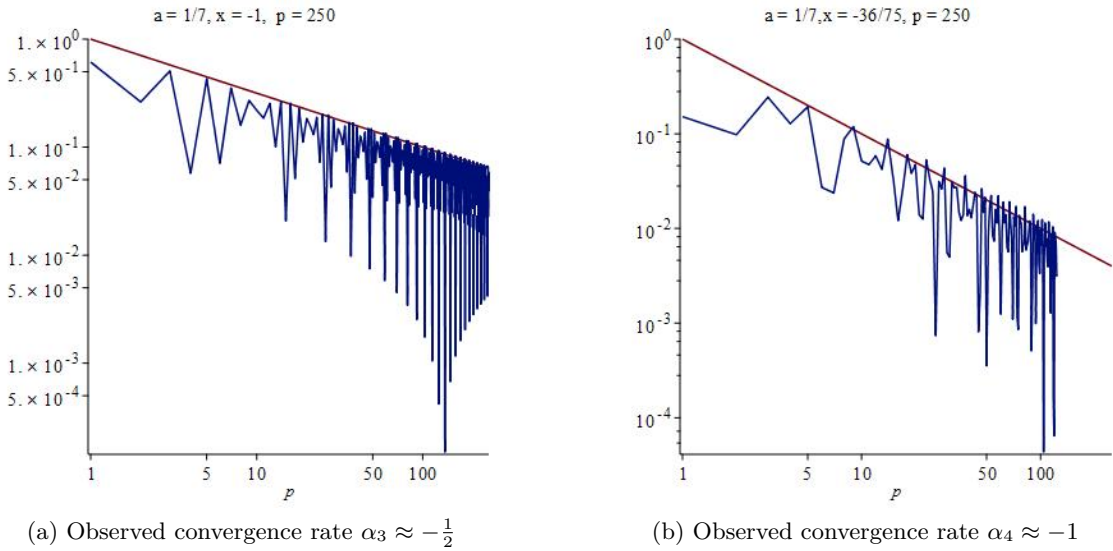


Figure 13: Pointwise error convergence rates of the Legendre polynomial expansion of function (12) with $a = \frac{1}{7}$ at points -1 (12b) and $\frac{36}{75}$ (13b)

Additionally, there was a noticeable difference in the behavior of the error e at two of the examined points. At point x_1 , we have $e_f(p, x) \approx \frac{1}{\sqrt{p}}$ while at point x_2 , we observe $e_f(p, x) \approx \frac{1}{p}$.

In order to understand the differences in behavior of error convergence rates, first we introduce the definition of the asymptotic rate.

Definition 2. $e_f(p, x)$ has asymptotic rate α for the function f if and only if

$$0 < \limsup e_f(p, x)p^\alpha = C_f(x) < \infty$$

By definition 2, the asymptotic rate is bounded above. Further experiments with different a and x revealed interesting results on the asymptotic and pre-asymptotic rates of convergence, and on the characteristic behavior of the error $e_f(p, x)$. Two rates of convergence were observed in the error approximation of a function $f(x)$ with a point of singularity $a = \frac{1}{2}$ at $x = -\frac{36}{75}$ with p up to 10^4 . The first, pre-asymptotic rate of convergence, was observed for p up to approximately 10^2 and it was slightly slower than the asymptotic rate at $p > 10^2$.

These and further results were then summarized in the following conjecture, where a corresponding theorem includes multiplicative $\lg p$ terms in the error estimates. This is made more precise below in the context of general Jacobi expansions.

Conjecture 2. Let $f(x) \in \Phi$ and $\mu > -1$

1. Let $x \in (-1, a) \cup (a, 1)$. Then $e_f(p, x) \leq C(x)p^{-\alpha}$, $\alpha = (\mu + 1)$, $C(x)$ is independent of p . The estimate is optimal in the sense that the estimate is not true for any $\alpha^* > (\mu + 1)$.
2. We have $C(1 + \xi) \leq D\xi^{-\rho}$, $C(1 - \xi) \leq D\xi^{-\rho}$, $\rho = \frac{1}{4}$, $0 < \xi \leq \delta$, $0 < \delta$ with D and δ independent of p and the rate $\rho = \frac{1}{4}$ is optimal.
3. We have $C(a + \xi) \leq D\xi^{-\sigma}$, $C(a - \xi) \leq D\xi^{-\sigma}$, $\sigma = 1$, $0 < \xi \leq \delta$, with D and δ independent of p and the rate $\sigma = 1$ is optimal.
4. Let $x = \pm 1$. Then $e_f(p, \pm 1) \leq Cp^{-\alpha}$, $\alpha = (\mu + \frac{1}{2})$, C independent of p . The rate $\alpha = \mu + \frac{1}{2}$ is optimal. For $\mu < -\frac{1}{2}$ we have $e_f(p, \pm 1) \rightarrow \infty$, for $\mu = -\frac{1}{2}$ we have $e_f(p, \pm 1)$ bounded.
5. Let $x = a$. Then $e_f(p, a) \leq Cp^{-\alpha}$, $\alpha = \mu + 1$, if β even, $\alpha = \mu$, otherwise. The rate $\alpha = \mu + 1$, if β even, $\alpha = \mu$, otherwise, is optimal.

4 Research question

Babuška and Hakula addressed the class of functions Φ approximated by the Legendre polynomials. In this thesis, the objective is to numerically study a pointwise Jacobi convergence conjecture based on the theorem of Legendre error (Conjecture 2).

For the class Φ of functions, the Jacobi expansion is defined

$$J_p^{\alpha, \beta}(f)(x) = \sum_{k=1}^p a_k^{(\alpha, \beta)} P_k^{\alpha, \beta}(x) \quad (98)$$

has the error

$$e_f(p, x, \alpha, \beta) = |f(x) - J_p^{\alpha, \beta}(f)(x)|. \quad (99)$$

The goal of this thesis is to verify the following Theorem, while the formal proof will be provided.

4.1 Theorem on $e_f(p, x, \alpha, \beta)$ for $f(x) \in \Phi$, $\mu > -1$

For the functions $f(x) \in \Phi$, $\mu > -1$, the proof of theorem for the Legendre polynomials can be extended as

Theorem 3. *Let $f(x) \in \Phi$ and*

1. *Let $x \in (-1, a) \cup (a, 1)$. Then $e_f(p, x, \alpha, \beta) \leq C(x, \alpha, \beta) p^{-\gamma} \lg p$, $\gamma = (\mu + 1)$, $C(x, \alpha, \beta)$ is independent of p . The estimate is optimal in the sense that the estimate is not true for any $\bar{\gamma} > (\mu + 1)$.*
2. *Let $x = -1$. Then $e_f(p, -1, \alpha, \beta) \leq C(-1, \alpha, \beta) p^{-\gamma} \lg p$, $\gamma = (\mu - \beta + \frac{1}{2})$, $C(-1, \alpha, \beta)$ independent of p . The rate $\gamma = \mu - \beta + \frac{1}{2}$ is optimal.*
3. *Let $x = 1$. Then $e_f(p, 1, \alpha, \beta) \leq C(1, \alpha, \beta) p^{-\gamma} \lg p$, $\gamma = (\mu - \alpha + \frac{1}{2})$, $C(1, \alpha, \beta)$ independent of p . The rate $\gamma = \mu - \alpha + \frac{1}{2}$ is optimal.*
4. *Let $x = a$. Then $e_f(p, a, \alpha, \beta) \leq C(a, \alpha, \beta) p^{-\gamma} \lg p$, $\gamma = \mu + 1$, if μ even, $\gamma = \mu$, otherwise. The rate $\gamma = \mu + 1$, if μ even, $\gamma = \mu$, otherwise, is optimal.*
5. *We have $C(-1 + \xi, \alpha, \beta) \leq D(\alpha, \beta) \xi^{-\rho}$, $\rho = 2\beta + \frac{1}{4}$, $0 < \xi \leq \delta$, $0 < \delta$ with $D(\alpha, \beta)$ and δ independent of p and the rate $\rho = 2\beta + \frac{1}{4}$ is optimal.*
6. *We have $C(1 - \xi, \alpha, \beta) \leq D(\alpha, \beta) \xi^{-\rho}$, $\rho = 2\alpha + \frac{1}{4}$, $0 < \xi \leq \delta$, $0 < \delta$ with $D(\alpha, \beta)$ and δ independent of p and the rate $\rho = 2\alpha + \frac{1}{4}$ is optimal.*
7. *We have $C(a + \xi, \alpha, \beta) \leq D(\alpha, \beta) \xi^{-\sigma}$, $C(a - \xi, \alpha, \beta) \leq D(\alpha, \beta) \xi^{-\sigma}$, $\sigma = 1$, $0 < \xi \leq \delta$, with $D(\alpha, \beta)$ and δ independent of p and the rate $\sigma = 1$ is optimal.*

Conjecture 4. *The multiplicative term $\lg p$ can be omitted from Theorem 3.*

5 Method

5.1 Calculation of Jacobi Expansion Coefficients

Consider a class of functions Φ , then the spectral expansion of $f \in \Phi$ is defined as in (98) where $a_n^{(\alpha,\beta)}$ are the Jacobi expansion coefficients that have a general definition

$$a_p^{(\alpha,\beta)} = \frac{1}{(P_p^{(\alpha,\beta)}(x), P_p^{(\alpha,\beta)}(x))} \int_{-1}^1 w^{(\alpha,\beta)}(x) J_p^{\alpha,\beta}(f)(x) P_p^{(\alpha,\beta)}(x) dx \quad (100)$$

In our case, however, the function f is not continuous on the whole interval $[-1, 1]$. Thus, we divide $[-1, 1]$ into two subintervals $[-1, a)$ and $(a, 1]$ such that f is continuous on each. This gives us a modified definition of the Jacobi expansion coefficients

$$a_p^{(\alpha,\beta)} = \frac{1}{(P_p^{(\alpha,\beta)}(x), P_p^{(\alpha,\beta)}(x))} \left[\int_{-1}^a J_p^{\alpha,\beta}(f)(x) P_p^{(\alpha,\beta)}(x) w^{(\alpha,\beta)}(x) dx - \int_a^1 J_p^{\alpha,\beta}(f)(x) P_p^{(\alpha,\beta)}(x) w^{(\alpha,\beta)}(x) dx \right]$$

Observe, $a_p^{(\alpha,\beta)}$ cannot be solved analytically and one is required to perform numerical integration. For this matter, we chose the software *Maple 2016* which is very efficient for such calculations. In Maple 2016, we used the package **orthopoly**, which is the package of orthogonal polynomials and the function **evalf** for numerical integration.

As the power p of Jacobi polynomial grows, the calculation of $a_p^{(\alpha,\beta)}$ becomes more time consuming. For example, the calculation of $a_{10}^{(-\frac{1}{2}, -\frac{1}{2})}$ takes approximately 10 seconds while the calculation of $a_{500}^{(-\frac{1}{2}, -\frac{1}{2})}$ can take up to 240 seconds (see Figure 14). The cumulative time for calculating the Jacobi expansion coefficients, hence, can take up to several hours.

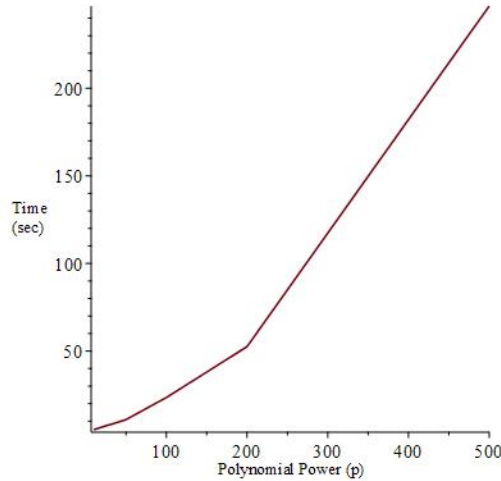


Figure 14: Computational time (in second) of $a_p^{(\alpha,\beta)}$ as a function of polynomial power, p .

In order to speed up the calculation of $a_p^{(\alpha,\beta)}$, we first simplified the algorithm by reducing the instances which require numerical integration. We define the inner product of

Jacobi polynomials as following

$$(P_p^{(\alpha,\beta)}(x), P_p^{(\alpha,\beta)}(x)) = \begin{cases} \int_{-1}^1 P_p^{(\alpha,\beta)}(x)P_p^{(\alpha,\beta)}(x)w^{(\alpha,\beta)}(x), & \text{if } p + \alpha + \beta + 1 \leq 0, \\ 2^{(\alpha+\beta+1)} \frac{\Gamma(p+\alpha+1)\Gamma(p+\beta+1)}{(2p+\alpha+\beta+1)n!\Gamma(p+\alpha+\beta+1)}, & \text{otherwise} \end{cases} \quad (101)$$

Note, since we have $\alpha > -1$ and $\beta > -1$, the first case can occur only when $p = 0$.

5.1.1 Parallelization of Numerical Integration

As mentioned earlier, the calculation of each $a_p^{(\alpha,\beta)}$ can take up to several minutes while the cumulative calculation time of p coefficients can take up to several hours. Observe, each $a_p^{(\alpha,\beta)}$ is independent of the previous $a_n^{(\alpha,\beta)}$ where $n = 0, \dots, p-1$, hence, the calculations of each are performed independently. This feature shows a potential for parallelization of the algorithm.

In order to execute the numerical integration in parallel, we have used the Maple 16 **Grid** package which supports multi-process parallel computation.

From Figure 14 we see that the computing time increases as p increases therefore, automatic distribution of the tasks among the processors will not provide the optimal improvement in the computation time. In other words, given a 4-processors machine, if one distributes the first $\frac{p}{4}$ coefficients to the first processor, second $\frac{p}{4}$ to the second and so on, the fourth processor will end up working by itself for a long time.

Alternative way to distribute the coefficients among available processors, is to first permute the coefficients, $a_p^{(\alpha,\beta)}$. The permutations can be performed by the combinatorial functions package **combinat** available on Maple 16. This way, each processor will compute coefficients of various powers and the computational time will decrease significantly.

5.2 Pointwise Error Approximation

In order to approximate the pointwise error, we first evaluated the function f at point x . Next, by definition (98), we get the Jacobi approximations at point x . Finally, using the definition (99), we calculated the pointwise error.

5.3 Approximation of Pointwise error convergence rates (Theorem 3 (1-4))

The convergence rates of pointwise error were calculated by using the set of approximated points in Section (5.2). Given two points $e(m, x, \alpha, \beta)$ and $e(n, x, \alpha, \beta)$, we can then express them in terms of the polynomial power and the coefficient $C(x, \alpha, \beta)$, see (102).

$$e(m, x, \alpha, \beta) = C(x, \alpha, \beta)m^r \quad (102)$$

where r is the convergence rate and m and n is the polynomial powers. Then, we can express the rate as follows.

$$r = \frac{\ln(e(m, x, \alpha, \beta)) - \ln(e(n, x, \alpha, \beta))}{\ln(m) - \ln(n)} \quad (103)$$

By Theorem 3, $C(x, \alpha, \beta)$ is independent of the polynomial power hence, it cancels out from the rate equation. Observed convergence rate is then the minimum rate between the selected sets of points.

5.4 Approximation of the Constants (Theorem 3 (5-7))

In order to approximate the constants defined in Theorem 3, we begin by assuming the convergence rate. For example, we assume that the pointwise error convergence rate at point $x = -1$ is $\gamma = \beta + \mu + \frac{1}{2}$ (Theorem 3 (2)). Then, for each point $x \pm \xi$ where $x \in \{-1, a, 1\}$ and $\xi \in \{0.1, 0.01, 0.001\}$, we compute the pointwise error. Finally, we can calculate the coefficients $C(x \pm \xi, \alpha, \beta)$ and their growth rates.

6 Results

6.1 Example 1 : $\mu \in \{0, 2, 2.5\}$, $a = \frac{1}{2}$, $\alpha = \beta = -\frac{1}{2}$

6.1.1 Verification of Theorem 3 (1-4)

For our first example, we considered three functions of class Φ corresponding to $\mu = 0, 2$, and 2.5 . We calculated the pointwise error convergence rates at four different points $x \in \{-1, \frac{1}{10}, a, 1\}$. Our findings correspond with Theorem 3 (1 – 4) for all μ . Summarized results can be found in Table 2 where γ and γ^* represent the expected and observed pointwise convergence rates, respectively. The results are also illustrated in Figures 15,17, and 23

x	Expected γ	$\mu = 0$		$\mu = 2$		$\mu = 2.5$	
		γ	γ^*	γ	γ^*	γ	γ^*
-1	$\mu - \beta + \frac{1}{2}$	1	0.9989	3	2.9969	3.5	3.4822
1	$\mu - \alpha + \frac{1}{2}$	1	0.9999	3	2.9979	3.5	3.4962
a	$\mu + 1$ or μ	0	0.0038	3	2.9872	2.5	2.4968
$\frac{1}{10}$	$\mu + 1$	1	1	3	3	3.5	3.5

Table 2: Expected and calculated pointwise error convergence rates of Jacobi polynomial approximation of a class Φ function with $\mu = 0, 2, 2.5$ at two boundary points $\{-1, 1\}$, one point of singularity a , and one interior point $\frac{1}{10}$ where $\alpha = \beta = -\frac{1}{2}$

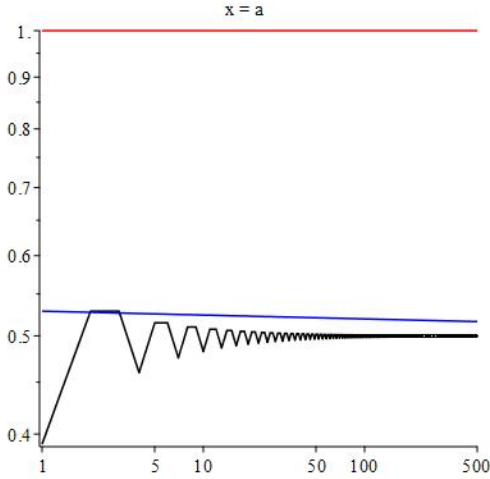
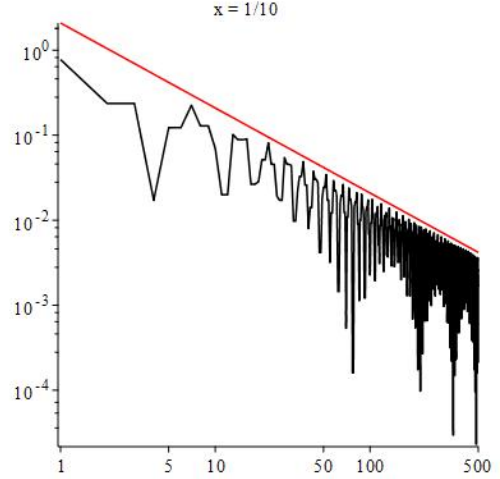
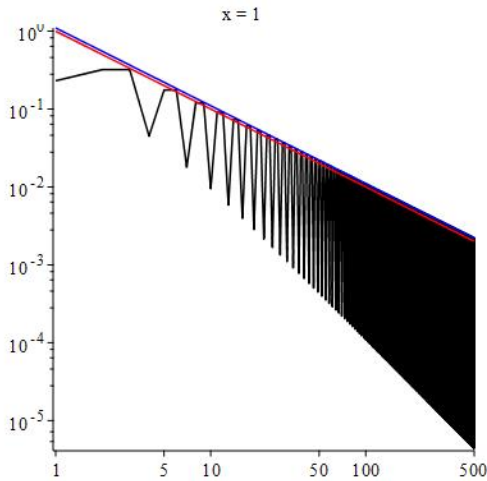
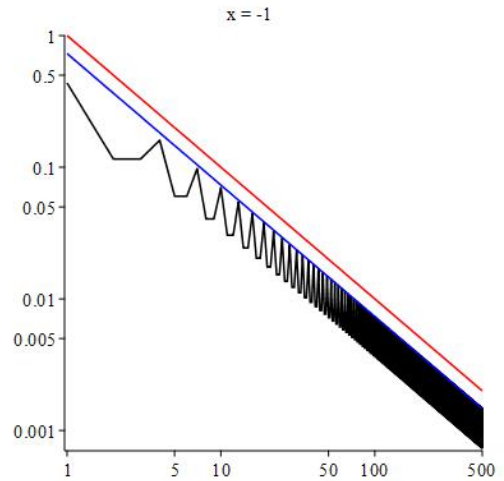
(a) $\gamma = 0, \gamma^* \approx 0.003823473681646$ (b) $\gamma = 1, \gamma^* \approx 1.$ (c) $\gamma = 1, \gamma^* \approx 0.99998367546827$ (d) $\gamma = 1, \gamma^* \approx 0.99898954889877$

Figure 15: Pointwise error (black), expected γ and calculated pointwise γ^* error convergence rate (red and blue, respectively) for $\mu = 0, a = \frac{1}{2}, \alpha = \beta = -\frac{1}{2}, p = 500$

From Figure 15, we see that pointwise error convergence rates are equivalent at all points on the interval $[-1, 1]$, except at point of singularity a . The observed convergence rate at point a can be explained by the behavior of (95). For $\mu = 0, a$ serves as a boundary point of second interval. In other words, we have a function

$$f(x) = \begin{cases} 0, & \text{for } -1 \leq x < a, \\ 1, & \text{for } a \leq x \leq 1, \end{cases} \quad (104)$$

Now, instead of (104), we take a different step function where a is a clear point of singularity,

such as:

$$f(x) = \begin{cases} -1, & \text{for } -1 \leq x < a, \\ 0, & \text{for } x = a, \\ 1, & \text{for } a < x \leq 1 \end{cases} \quad (105)$$

Then, we observe pointwise error convergence rate of $\gamma^* = 0.9947$ at point a , see Figure 16

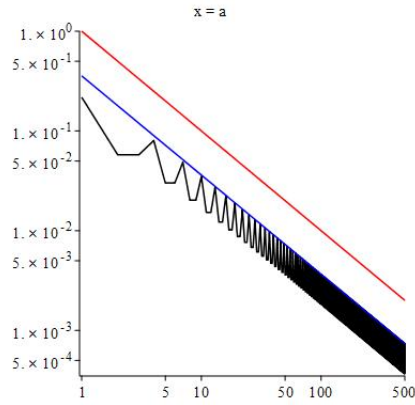
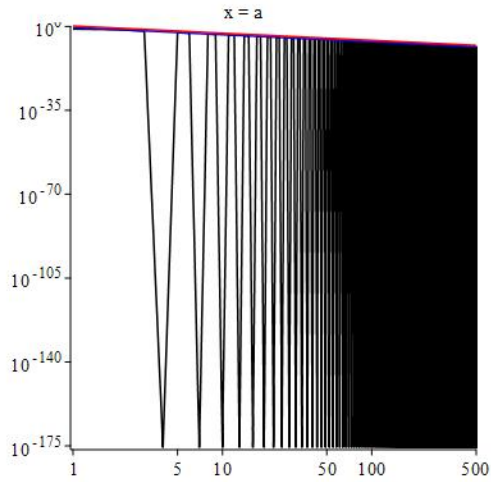
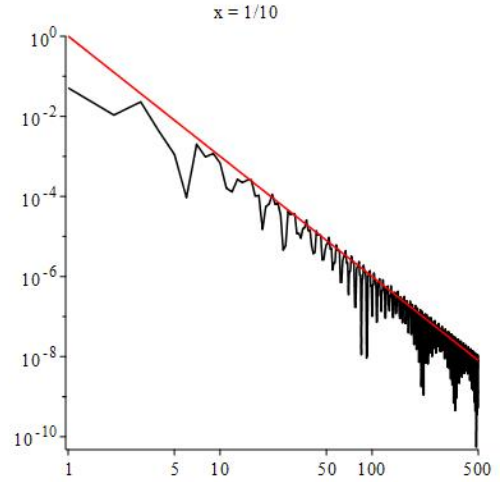
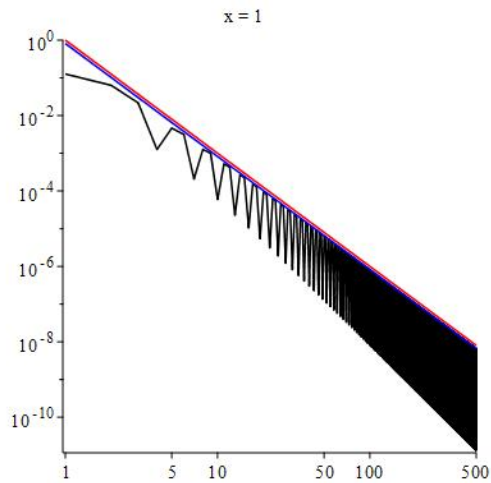
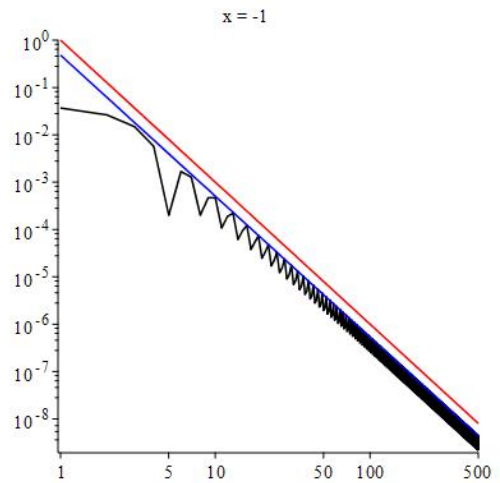


Figure 16: Pointwise error convergence at point a considering a step function (105)

In addition to the convergence rates, one can observe a repeating pattern of the pointwise error at points a , -1 , and 1 .

(a) $\gamma = 3, \gamma^* \approx 2.98724136$ (b) $\gamma = 3, \gamma^* \approx 3$ (c) $\gamma = 3, \gamma^* \approx 2.99796735$ (d) $\gamma = 3, \gamma^* \approx 2.99696662$ Figure 17: $\mu = 2, a = 0.5, \alpha = \beta = -0.5, p = 500$

In Figure 17, we see the pointwise error and its calculated convergence rate of Jacobi polynomial approximation of (95) with $\mu = 2$. Here we observe an equivalent convergence rates at all points on the interval $[-1, 1]$ and a recurring pattern of pointwise error at points $-1, 1,$ and a .

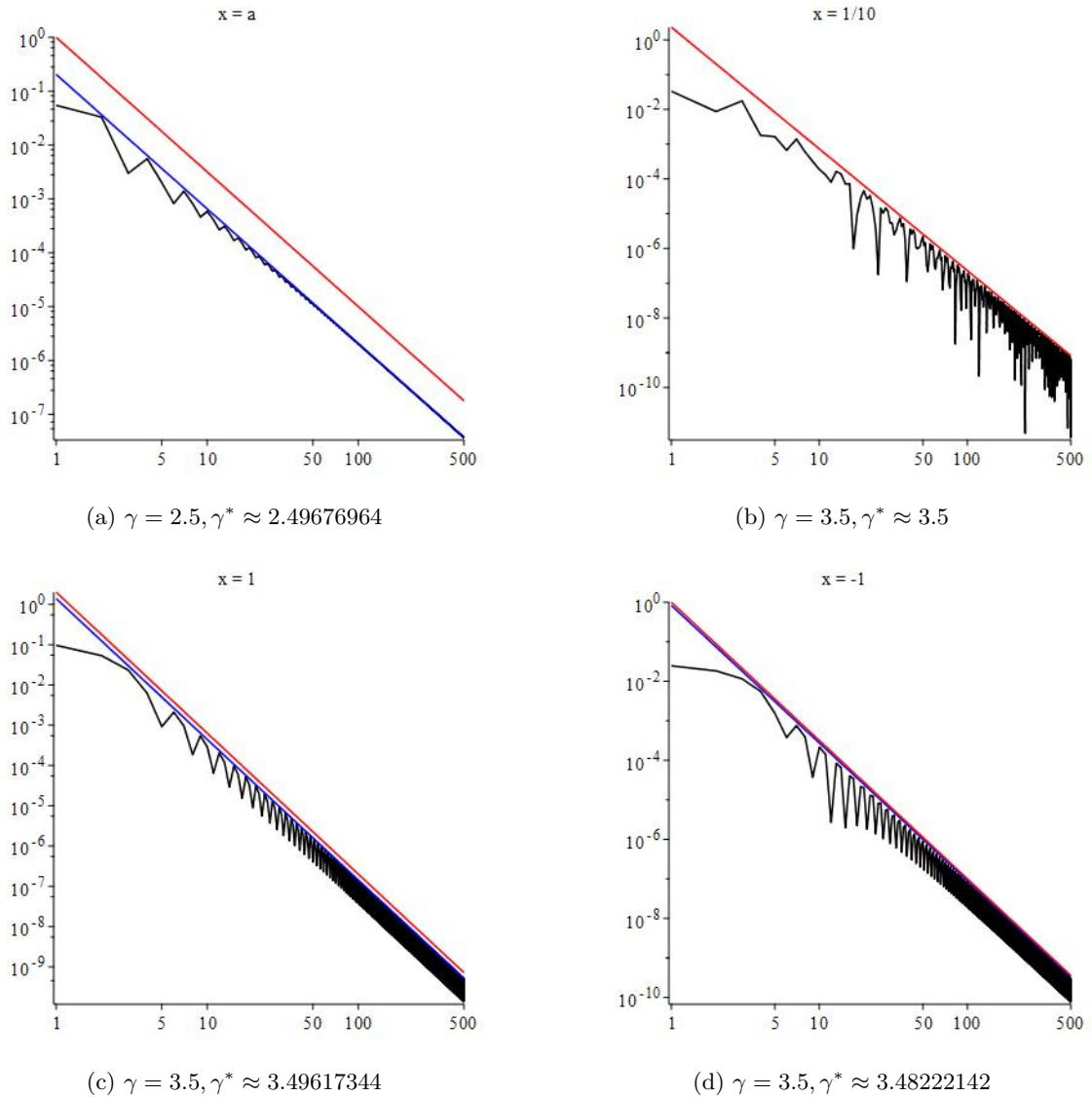


Figure 18: $\mu = 2.5, a = 0.5, \alpha = \beta = -0.5, p = 500$

Finally, the pointwise error of Jacobi polynomial approximation of (95) with $\mu = 2.5$ is illustrated in Figure 18 for four different points on interval $[-1, 1]$. We observe same convergence rates at points $-1, \frac{1}{10},$ and 1 . Unlike in two previous examples, the pointwise error convergence rate at point of singularity a is slower. Theorem 3 has described this behavior in (4), where the optimal convergence rate $\gamma = \mu$ is optimal for $\mu \notin \mathbb{Z}$.

In each of the first three examples, we observed a similar pointwise convergence rates for the boundary and interior points. The pointwise error convergence was different at point a , where for $\mu = 0$ we observed no convergence at all, and for $\mu = 2.5$ the convergence rate was slower from other points. We also saw that with the increase of degree μ of (95), the pointwise error convergence rate increased, as well.

6.1.2 Verification of Theorem 3 (5-7)

Consider $P_p^{\alpha,\beta}$ where $\alpha = \beta = -\frac{1}{2}$ up to degree $p = 500$ and (95) with $\mu = 0, 2$, and 2.5 . Recall, for our constant calculations we assume the optimal pointwise error convergence rates stated in Theorem 3 (see Section 5.4). Based on our calculations in Section 6.1.1, the convergence rates for polynomial orders $p = 0, \dots, 500$ do not reach optimal rates stated in the Theorem. Thus, the assumption of optimal rates leads to errors in calculations and consequently, we cannot compute the constants with high accuracy.

In this section, however, we are more interested in the growth rates of constants as we approach boundary and singularity points. Although, low accuracy of constants reduces the accuracy of growth rates, we can still observe the behavior of constants at the neighborhood of discontinuity points. For our calculations, we considered the boundary points $\{-1, 1\}$ and the point of singularity a . Since points $x \pm \xi$ are all in the interior, we refer to Theorem 3(1) where we find the expected convergence rates γ . Using γ we then calculate the constants for points $\{x \pm \xi, x \pm \xi^2, x \pm \xi^3\}$. The optimal growth rates based on Theorem 3 are:

- At point -1 , $\rho = 2\beta + \frac{1}{4} = -\frac{3}{4}$
- At point a , $\sigma = 1$
- At point 1 , $\rho = 2\alpha + \frac{1}{4} = -\frac{3}{4}$

Observe, according to Theorem 3, constants are independent of the polynomial order, p , and of the function (95) order, μ . Thus, increasing or decreasing either will not be reflected in the growth rates of the constants. In order to verify the statement above, we calculated constants for $\mu = 0, 2$, and 2.5 .

In tables 3, 4, and 5 below, we summarized the obtained constants for each of the points. For each μ and each point $\{-1, a, 1\}$, we calculated two rates

- σ_1/ρ_1 - growth rate from $x \pm \xi$ to $x \pm \xi^2$
- σ_2/ρ_2 - growth rate from $x \pm \xi$ to $x \pm \xi^3$

$x + \xi$				$x - \xi$			
$-1 + \xi$	0.766	$a + \xi$	5.5115	$a - \xi$	6.029	$1 - \xi$	1.378
$-1 + \xi^2$	0.736	$a + \xi^2$	42.997	$a - \xi^2$	41.420	$1 - \xi^2$	1.125
$-1 + \xi^3$	0.734	$a + \xi^3$	203.811	$a - \xi^3$	319.493	$1 - \xi^3$	1.105

Table 3: Calculated constants $C(x, \alpha, \beta)$ near boundary and singularity points for $\alpha = \beta = -\frac{1}{2}$ and $\mu = 0$

All calculated growth rates one can find in Figures (19), (20) and (21).

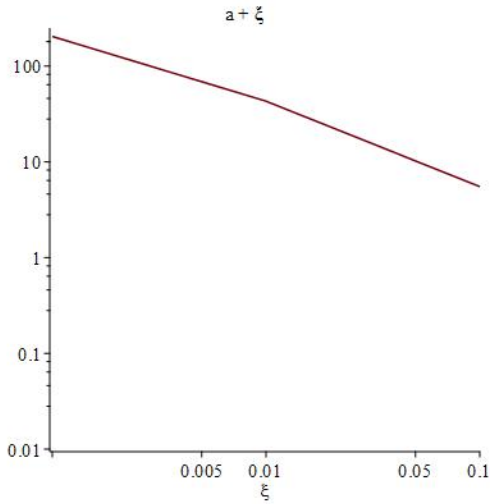
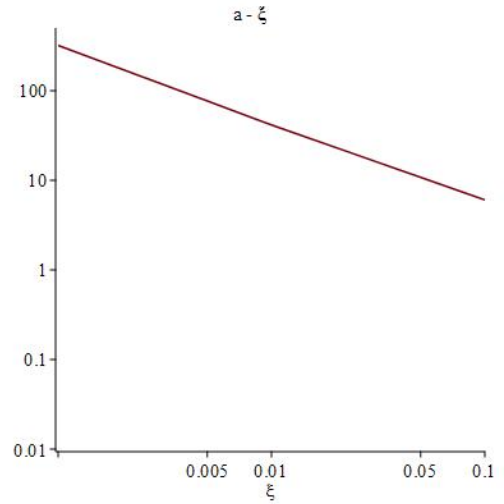
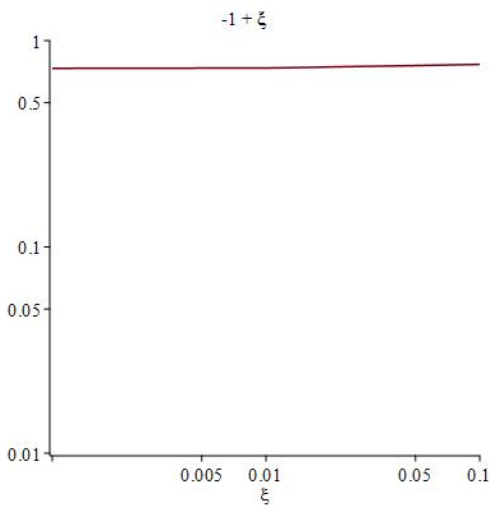
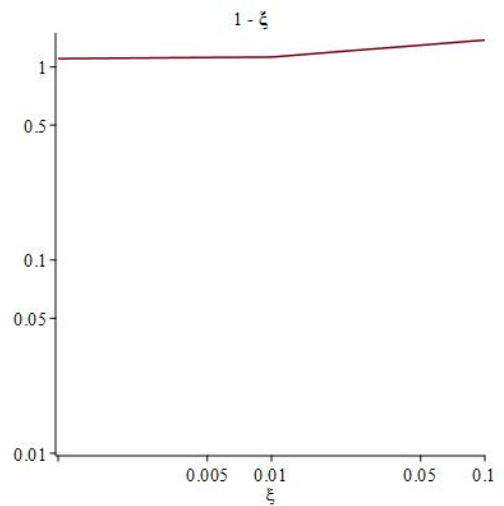
(a) $x = a + \xi$, $\sigma_1 = 0.8922$, $\sigma_2 = 0.7840$ (b) $x = a - \xi$, $\sigma_1 = 0.8370$, $\sigma_2 = 0.8621$ (c) $x = -1 + \xi$, $\rho_1 = -0.0174$, $\rho_2 = -0.0095$ (d) $x = 1 - \xi$, $\rho_1 = -0.0881$, $\rho_2 = -0.0480$

Figure 19: Growth rates of constants $C(x, \alpha, \beta)$ for function (95) with $\mu = 0$ as we move closer to the boundary and singularity points. (σ_1/ρ_1 - growth rate from $x \pm \xi$ to $x \pm \xi^2$, σ_2/ρ_2 - growth rate from $x \pm \xi$ to $x \pm \xi^3$)

First we examined the constants obtained from Jacobi polynomials approximation of a step function (95) with $\mu = 0$. At the neighborhood of boundary points -1 and 1 , we observe growth rates very close to 0, meaning that the constants do not grow as we move closer to the boundaries. At the neighborhood of singularity point, the constants growth rate σ^* is significantly higher than at the boundaries but $\sigma^* < \sigma = 1$, where σ is the optimal constant growth rate (Figure 19).

$x + \xi$				$x - \xi$			
$-1 + \xi$	0.574	$a + \xi$	4.124	$a - \xi$	4.494	$1 - \xi$	1.031
$-1 + \xi^2$	0.551	$a + \xi^2$	22.222	$a - \xi^2$	20.441	$1 - \xi^2$	0.842
$-1 + \xi^3$	0.549	$a + \xi^3$	82.989	$a - \xi^3$	82.623	$1 - \xi^3$	0.827

Table 4: Calculated constants $C(x, \alpha, \beta)$ near boundary and singularity points for $\alpha = \beta = -\frac{1}{2}$ and $\mu = 2$

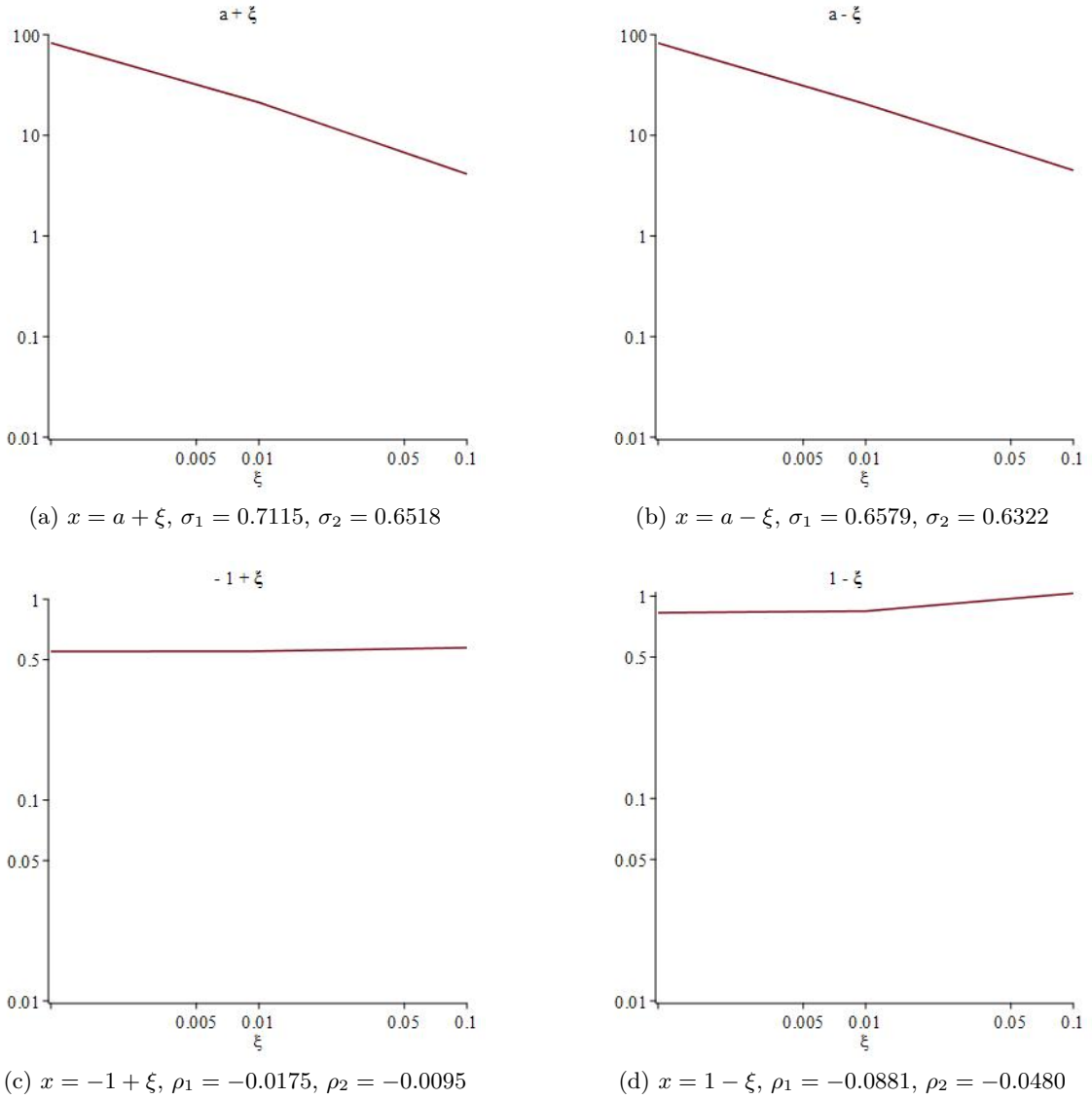


Figure 20: Growth of constants $C(x, \alpha, \beta)$ for function (95) with $\mu = 2$ as we move closer to the boundary and singularity points. (σ_1/ρ_1 - growth rate from $x \pm \xi$ to $x \pm \xi^2$, σ_2/ρ_2 - growth rate from $x \pm \xi$ to $x \pm \xi^3$)

We proceeded examining the constants and their growth rates for (95) with $\mu = 2$. Since constants $C(x, \alpha, \beta)$ are not a function of μ , we do not expect any significant difference in their behavior from previous example with $\mu = 0$. Indeed, the growth rates of constants

observed at the neighborhood of -1 and 1 are close to 0 (Figure 20). At the neighborhood of singularity point, we observed higher growth rates of the constants, $\sigma_{a+\xi}^* = 0.6518$ and $\sigma_{a-\xi}^* = 0.6322$, however, neither have reached the optimal growth rate of $\sigma = 1$.

$x + \xi$				$x - \xi$			
$-1 + \xi$	0.856	$a + \xi$	4.677	$a - \xi$	6.8882	$1 - \xi$	1.696
$-1 + \xi^2$	0.822	$a + \xi^2$	10.527	$a - \xi^2$	52.136	$1 - \xi^2$	1.437
$-1 + \xi^3$	0.819	$a + \xi^3$	12.470	$a - \xi^3$	133.617	$1 - \xi^3$	1.422

Table 5: Calculated constants $C(x, \alpha, \beta)$ near boundary and singularity points for $\alpha = \beta = -\frac{1}{2}$ and $\mu = 2.5$

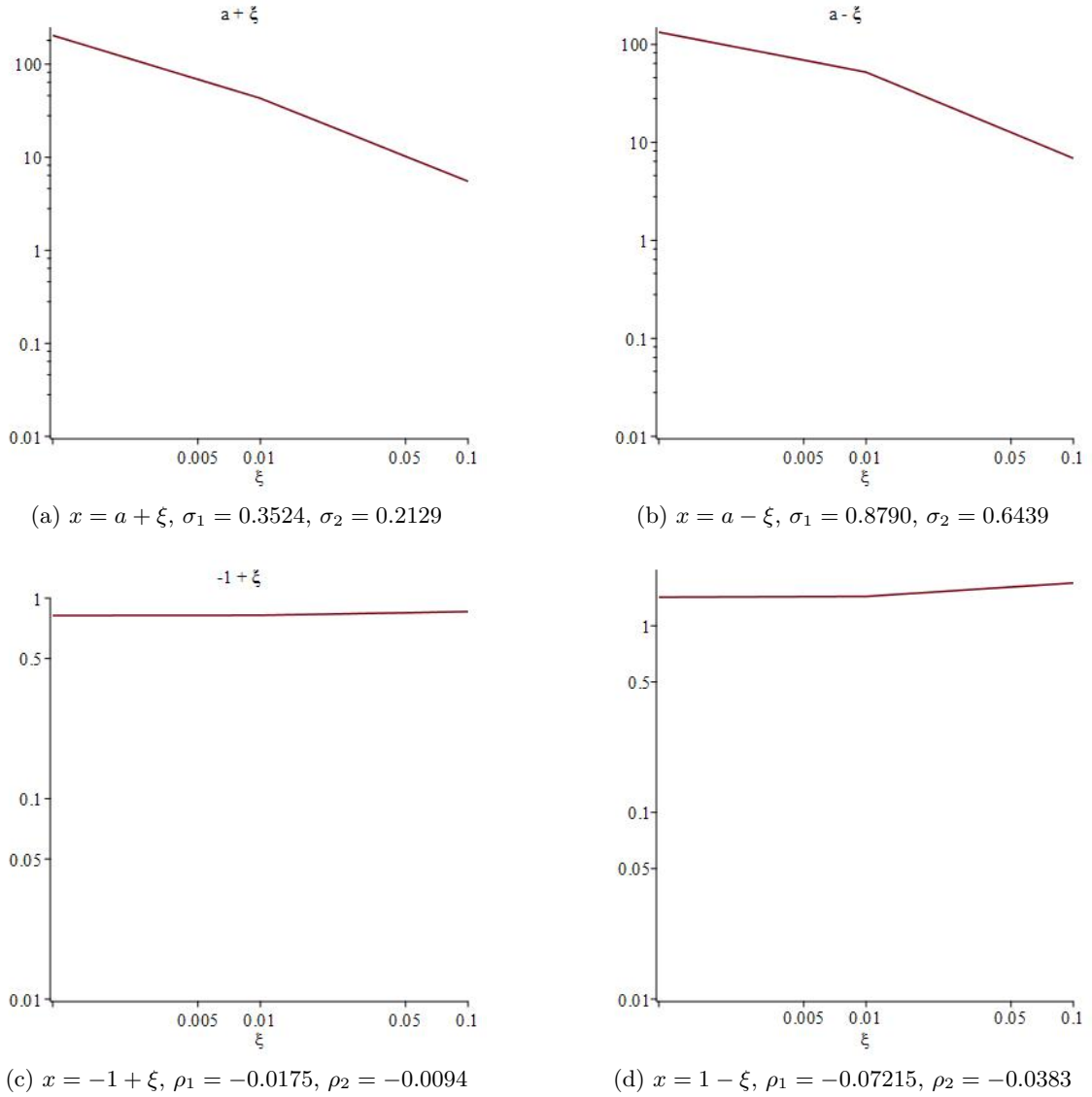


Figure 21: Growth of constants $C(x, \alpha, \beta)$ for function (95) with $\mu = 2.5$ as we move closer to the boundary and singularity points. (σ_1/ρ_1 - growth rate from $x \pm \xi$ to $x \pm \xi^2$, σ_2/ρ_2 - growth rate from $x \pm \xi$ to $x \pm \xi^3$)

Lastly we have considered (95) with $\mu = 2.5$. Our observation of constants' behavior and growth rates do not reveal anything different from the previous two examples. Similarly to the past instances, the growth rates at the boundaries were very small. Although the observed growth rates near the point of singularity are higher, neither surpass the optimal growth rates. One can find the calculated growth rates in Figure 21.

In Table 6, we summarized the computed growth rates at the boundary points $\{-1, 1\}$. One can directly note that constants and their growth rates are independent of μ and p . At point of singularity a , we observed a slight fluctuation in growth rates for different μ . This, however, might be a result of our pointwise error assumption.

	$\mu = 0$		$\mu = 2$		$\mu = 2.5$	
	σ_1 / ρ_1	σ_2 / ρ_2	σ_1 / ρ_1	σ_2 / ρ_2	σ_1 / ρ_1	σ_2 / ρ_2
$-1 + \xi$	-0.017418	-0.009494	-0.017477	-0.009499	-0.0175422	-0.009434
$1 - \xi$	-0.088070	-0.047991	-0.088138	-0.047984	-0.072151	-0.038305

Table 6: Comparison of constants' growth rates as we approached boundary points $\{1, 1\}$ for different μ .

Our study of $C(x, \alpha, \beta)$ at points $x \in \{-1, a, 1\}$ and $\xi \in \{0.1, 0.01, 0.001\}$ has shown that the constants are, indeed, independent of polynomial order p and the order of (95) μ . Moreover, as we moved closer to the singularity point a by reducing ξ , the constants grew. All observed growth rates were well below the optimal growth rates discussed in Theorem 3 (5-7).

6.2 Example 2 : $\mu = 0, a = \frac{1}{2}, \alpha = 1, \beta = 2$

For our second test of pointwise error convergence rate, we considered same points $\{-1, \frac{1}{10}, a, 1\}$ where the point of singularity remained, $a = \frac{1}{2}$. Here, we chose $\mu = 0$, $\alpha = 1$, and $\beta = 2$. Moreover, in this and next example, we examined the pointwise error of Jacobi polynomial approximation of (105), instead of (95). The behavior of pointwise error is similar across all points between the two functions, except at point a , where we expect to see no convergence at all given (95) with $\mu = 0$. Hence, with (105), we expect to see pointwise error convergence rates at all points.

In Table 7, one can find the calculated pointwise error convergence rates which are also illustrated in Figure 22. Again, γ and γ^* correspond to the expected and calculated convergence rates.

x	Expected γ	γ	γ^*
-1	$\mu - \beta + \frac{1}{2}$	$-\frac{3}{2}$	-1.4925
1	$\mu - \alpha + \frac{1}{2}$	$-\frac{1}{2}$	-0.4979
a	1	1	0.9951
$\frac{1}{10}$	$\mu + 1$	1	≈ 1

Table 7: Expected and calculated pointwise error convergence rates of Jacobi polynomial approximation of function (105) at two boundary points $\{-1, 1\}$, one point of singularity a , and one interior point $\frac{1}{10}$ where $\alpha = 1$ and $\beta = 2$

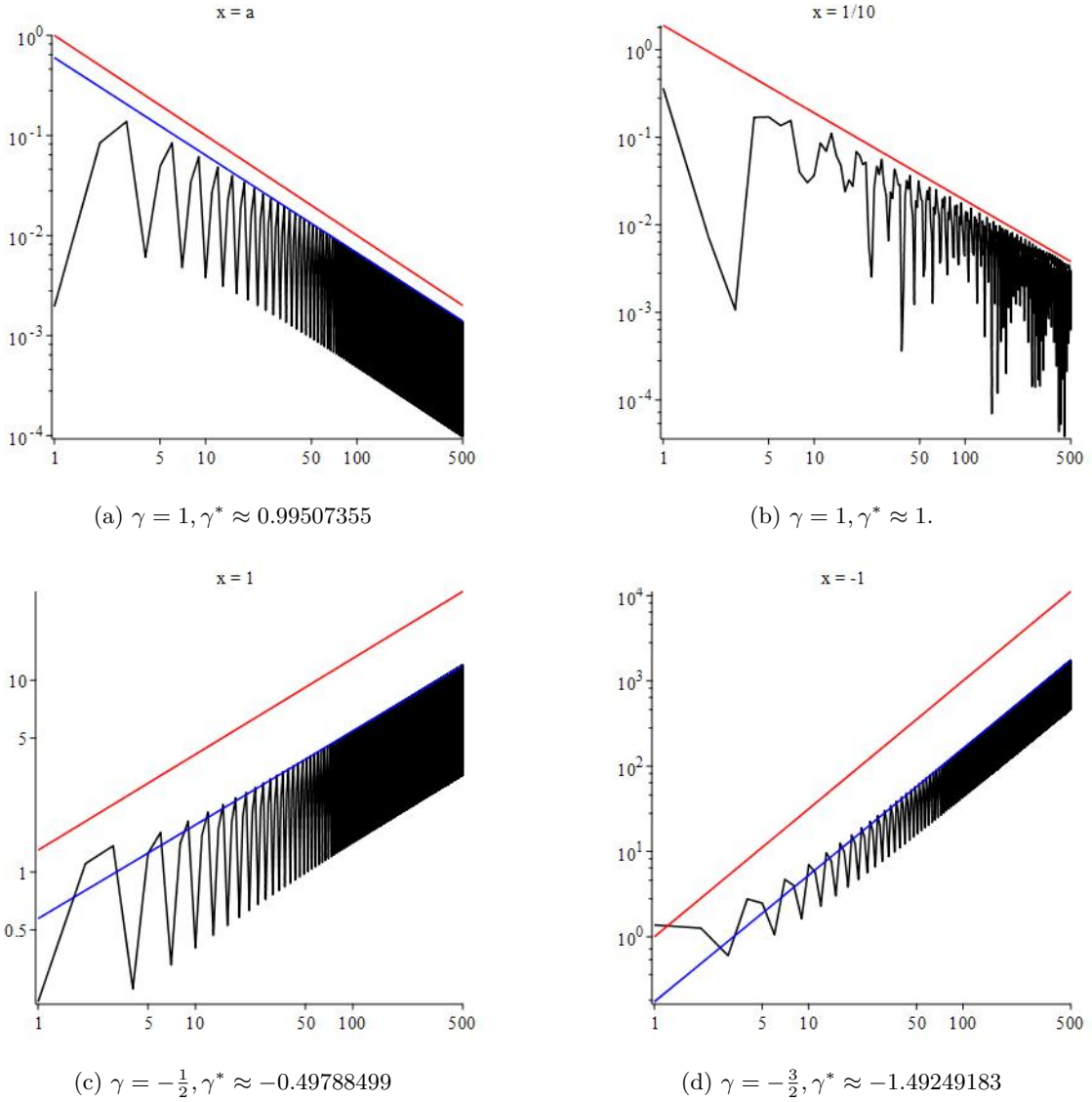


Figure 22: Pointwise error (black), expected γ and calculated pointwise γ^* error convergence rate (red and blue, respectively) in Jacobi polynomial approximation of function (105), where $a = \frac{1}{2}$, $\alpha = 1$ and $\beta = 2$, $p = 500$

The results of this example correspond with the expected pointwise error convergence rates in Theorem 3 (1-4). At the boundary points -1 and 1 we observe a divergence of the pointwise error which means that, unlike at interior and singularity points, increase in the power of $P_p^{\alpha,\beta}$ leads to increase in the error at these points. Furthermore, we again observe a recurring pattern of pointwise error at points -1 , a , and 1 .

6.3 Example 3 : $\mu = 0$, $a = \frac{1}{2}$, $\alpha = -\frac{1}{2}$, $\beta = \frac{1}{2}$

In our previous examples, we have considered α and β , either both positive or both negative. In our third and last example, we examine the behavior of pointwise error when $\alpha = -\frac{1}{2}$ and $\beta = \frac{1}{2}$. Note, due to insufficient machine memory, the computation of pointwise error was possible only up to $p = 100$.

The pointwise error convergence rate for this example are summarized in Table 8

x	Expected γ	γ	γ^*
-1	$\mu - \beta + \frac{1}{2}$	0	0
1	$\mu - \alpha + \frac{1}{2}$	1	0.9893
a	1	1	0.9947
$\frac{1}{10}$	$\mu + 1$	1	≈ 1

Table 8: Expected and calculated pointwise error convergence rates of Jacobi polynomial approximation of function (105) at two boundary points $\{-1, 1\}$, one point of singularity a , and one interior point $\frac{1}{10}$ where $\alpha = -\frac{1}{2}$ and $\beta = \frac{1}{2}$

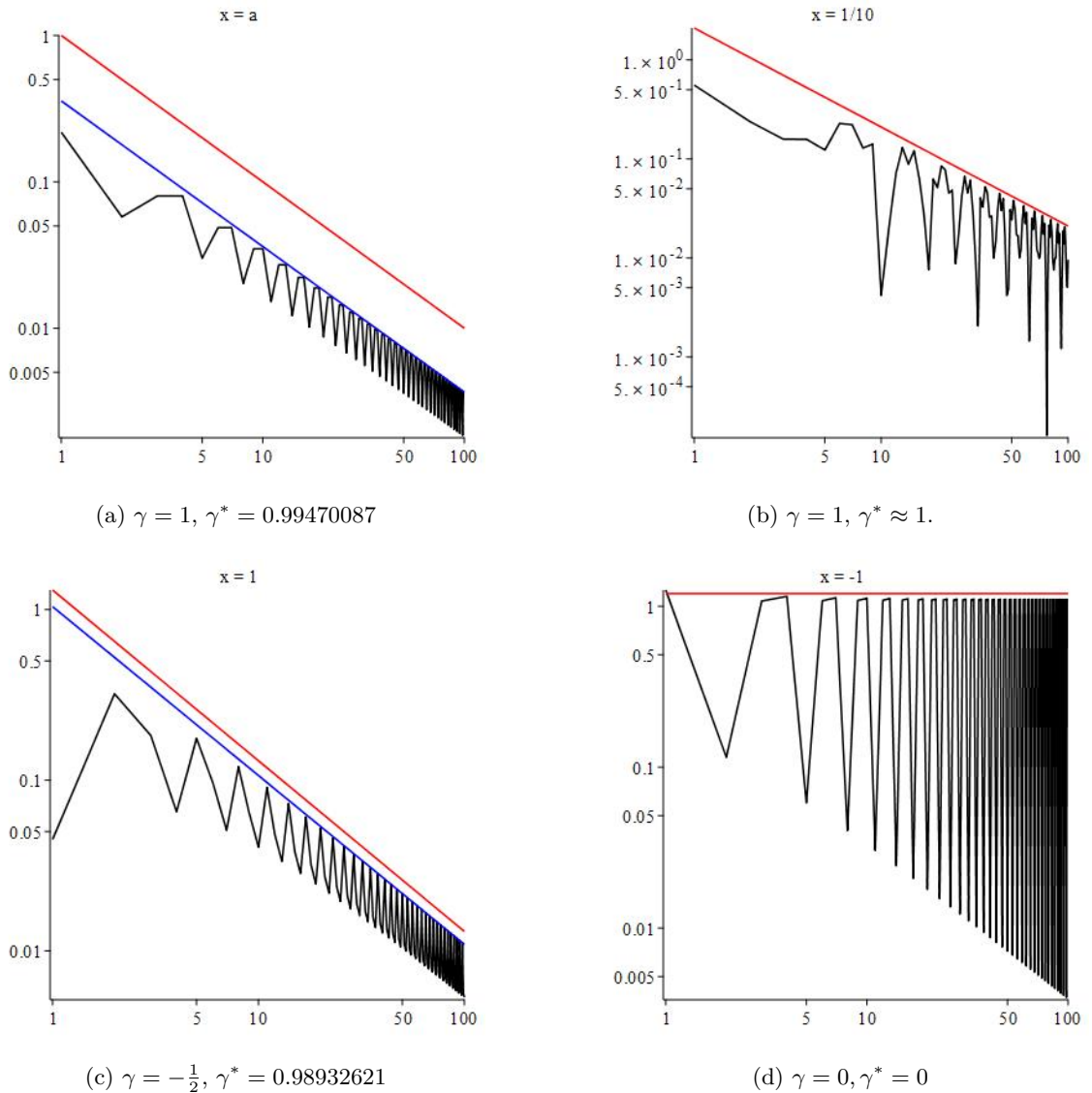


Figure 23: Pointwise error (black), expected γ and calculated pointwise γ^* error convergence rate (red and blue, respectively) in Jacobi polynomial approximation of function (105), where $a = \frac{1}{2}$, $\alpha = -\frac{1}{2}$ and $\beta = \frac{1}{2}$, $p = 100$

The observed pointwise convergence rates agree with the expected rates of Theorem 3. At a left boundary point -1 , we observe a convergence rate of 0, hence the pointwise error remains the same as we increase the degree of polynomial p . The pointwise error convergence rate was equivalent at the rest of points.

7 Discussion

Results of our numerical tests correspond with the statements of Theorem 3. The pointwise error convergence rate of Jacobi polynomial approximation of a class Φ functions depends on several things described in Theorem 3. First of all, it depends on the point of interest x , in particular, whether it is a boundary, singularity, or interior point. Secondly, the pointwise error convergence rate depends on α and β . Lastly, of course, it depends on the degree μ of (95).

With Theorem 3 and our results, we can now predict the pointwise error convergence rates of all polynomials that are special cases of Jacobi polynomials. In our first example, we have considered $P_p^{(\alpha,\beta)}$ with $\alpha = \beta = -\frac{1}{2}$, which is usually referred to as Chebyshev polynomials of the first kind (Section 2.2.3). Now, considering Chebyshev polynomials of the second kind (Section 2.2.4), where $\alpha = \beta = \frac{1}{2}$, and a class Φ function with $\mu = 0$, we expect no convergence at points a , 1 , and -1 . The only pointwise error convergence greater than 0 is expected on the interior of $[-1, 1]$.

In our experiments, we observed the pointwise error convergence rates of Jacobi polynomials up to power 500. Although we have obtained expected convergence rates, it is necessary to perform these experiments for higher powers of p . Higher polynomial orders are crucial for accurate computations of constants and their growth rates. Additionally, it is necessary to perform the calculation of Example 6.3 with $p > 100$, although this might require a different algorithm and possibly a more powerful machine.

To further emphasize the importance of higher polynomial orders for study of pointwise error convergence rates, recall the study of Legendre polynomials [Babuška and Hakula, 2016]. In their experiments, Babuška and Hakula measured the pointwise error convergence rates of Legendre polynomials raised up to power $p = 10^4$ and, thus, were able to observe both asymptotic and pre-asymptotic rates. Due to low polynomial order in all of our experiments, we were not able to observe the existence of pre-asymptotic and asymptotic rates of convergence.

References

- I. Babuška and D. Q. Guo. The $h - p$ version of the finite element method for domains with curved boundaries. *SIAM Journal on Numerical Analysis*, 25:837–861, 2006. URL <https://doi.org/10.1137/0725048>.
- I. Babuška and H. Hakula. Pointwise error estimate of the legendre expansion: The known and unknown features. *ask*, pages 0–0, 2016.
- I. Babuška, B.A.Szabo, and I. Katz. The p -version of the finite element method. *SIAM J. NUMER. ANAL.*, 18:515–545, 1981.
- N. K. Bary. *A Treatise on Trigonometric Series*, volume 2. Macmillan, 1964.
- P. Borwein and T. Erdelyi. *Polynomials and Polynomial Inequalities*, volume 1. Springer, 1995.
- E. Britannica.
- R. Courant and E. McShane. *Differential and Integral Calculus*, volume 1. John Wiley & Sons Inc, 2011. doi: 10.1002/9781118033241.
- D. Gottlieb and S. Orszag. *Numerical Analysis of Spectral Methods: Theory and Applications*. SIAM Society for Industrial and Applied Mathematics, 1977. URL <https://doi.org/10.1137/1.9781611970425>.
- D. Gottlieb and C. Shu. On the gibbs phenomenon and its resolution. *SIAM Rev.*, 39: 644–668, 1997.
- L. Guilbeau. The history of the solution of the cubic equation. *Mathematical Association of America*, 5:8–12, 1930.
- D. Jackson. The general theory of approximation by polynomials and trigonometric sums. *Bull. Amer. Math. Soc.*, 27:415–431, 1921.
- V. Lakshminarayanan and L. Varadharajan. *Special Functions for Optical Science*, volume 2. SPIE, 2015.
- C. Oddy. The gibbs phenomenon and its resolution. Master’s thesis, Linnaeus University, 2015.
- L. R. Scott. *Numerical Analysis*. Princeton University Press, 2011.
- J. Shen, T. Tang, and L. Wang. *Spectral Methods: Algorithms, Analysis, and Applications*. Springer, 2011.
- B. Szabo and I. Babuška. *Introduction to Finite Element Analysis: Formulation, Verification, and Validation*. Wiley, 2011.
- C. Szego. *Orthogonal Polynomials*. American Mathematical Society, 1939.
- L. B. Wahlbin. A comparison of the local behavior of spline L_2 -projections, Fourier series and Legendre series. In P. Grisvard, W. Wendland, and J. Whiteman, editors, *Singularities and Constructive Methods for Their Treatment*, volume 1121 of *Lecture Notes in Mathematics*, pages 319–346. Springer Berlin Heidelberg, 1985. ISBN 978-3-540-15219-4. doi: 10.1007/BFb0076279. URL <http://dx.doi.org/10.1007/BFb0076279>.

- H. Wang and S. Xiang. On the convergence rates of Legendre approximation. *Mathematics of Computation*, 81(278):861–877, 2012.
- C. J. Zarowski. *An Introduction to Numerical Analysis for Electrical and Computer Engineers*.
- D. Zwillinger. *CRC Standard Mathematical Tables and Formulae*. CRC Print, 32 edition, 2011.