Approximations in Bernstein Form

Peter Occil

This version of the document is dated 2025-09-30.

Peter Occil

1 Introduction

This page describes how to compute a polynomial in Bernstein form that comes close to a known function $f(\lambda)$ with a user-defined error tolerance, so that the polynomial's Bernstein coefficients will lie in the closed unit interval if f's values lie in that interval. The polynomial is often simpler to calculate than the original function f and can often be accurate enough for an application's purposes.

The goal of these approximations is to avoid introducing transcendental and trigonometric functions to the approximation method. (Therefore, although this page also discusses approximation by so-called *Chebyshev interpolants*, that method is relegated to the appendix.)

> **Notes:**
>
> 1. This page was originally developed as part of a section on *approximate Bernoulli factories*, or algorithms that toss heads with probability equal to a polynomial that comes close to a continuous function. However, the information in this page is of much broader interest than the approximate Bernoulli factory problem.
>
> 2. In practice, the level at which the function $f(\lambda)$ is known may vary:
>
>   1. $f(\lambda)$ may be known so completely that any property of f that is needed can be computed (for example, $f(\lambda)$ is given in a symbolic form such as $\sin(\lambda)/3$ or $\exp(-\lambda/4)$). Or…
>   2. f may be given as a "black box", but it's possible to find the exact value of $f(\lambda)$ for any λ (or at least any rational λ) in f's domain. Or…
>   3. Only the values of f at a finite number of points (such as equally spaced points) may be known.
>
>   In the last two cases, additional assumptions on f may have to be made in practice, such as upper bounds on f's first or second derivative, or whether f has a continuous r-th derivative for every r (see "Definitions"). If f does not meet those assumptions, the polynomial that approximates f will not necessarily achieve the desired accuracy.[1]

2 Contents

- **Introduction**
- **Contents**

---

[1]In case 3 in general, if f is analytic at every point on an interval, the "most stable" approximation occurs when the sample points are clustered at a quadratic rate toward the endpoints. Adcock, B., Platte, R.B., Shadrin, A., "**Optimal sampling rates for approximating analytic functions from pointwise samples**, *IMA Journal of Numerical Analysis* 39(3), July 2019. https://doi.org/10.1093/imanum/dry024

3   About This Document

This is an open-source document; for an updated version, see the source code[2] or its rendering on GitHub[3]. You can send comments on this document on the GitHub issues page[4], especially if you find any errors on this page.

My audience for this article is **computer programmers with mathematics knowledge, but little or no familiarity with calculus**.

4   Definitions

This section describes certain math terms used on this page for programmers to understand.

The *closed unit interval* (written as $[0, 1]$) means the set consisting of 0, 1, and every real number in between.

For definitions of *continuous*, *derivative*, *convex*, *concave*, *Hölder continuous*, and *Lipschitz continuous*, see the definitions section in "**Supplemental Notes for Bernoulli Factory Algorithms[5]**".

Any polynomial $P(\lambda)$ can be written in *Bernstein form* as—

$$P(\lambda) = a[0]p_{n,0}(\lambda) + ... + a[n]p_{n,n}(\lambda), p_{n,k}(\lambda) = \binom{n}{k}\lambda^k(1-\lambda)^{n-k},$$

where n is the polynomial's *degree* and $a[0]$, $a[1]$, ..., $a[n]$ are its n plus one *Bernstein coefficients* (which this document may simply call *coefficients* if the meaning is obvious from the context).[6]

---

[2]https://github.com/peteroupc/peteroupc.github.io/raw/master/bernapprox.md

[3]https://github.com/peteroupc/peteroupc.github.io/blob/master/bernapprox.md

[4]https://github.com/peteroupc/peteroupc.github.io/issues

[5]https://peteroupc.github.io/bernsupp.html#Definitions

[6]choose(n, k) = (1*2*3*...*n)/((1*...*k)*(1*...*$(n-k)$)) = $n!/(k! * (n-k)!)$ = $\binom{n}{k}$ is a *binomial coefficient*, or the number of ways to choose k out of n labeled items. It can be calculated, for example, by calculating $i/(n-i+1)$ for each integer i satisfying $n-k+1 \le i \le n$, then multiplying the results (Yannis Manolopoulos. 2002. "Binomial coefficient computation: recursion or iteration?", SIGCSE Bull. 34, 4 (December 2002), 65–67. DOI: **https://doi.org/10.1145/820127.820168**). For every $m>0$, choose(m, 0) = choose(m, m) = 1 and choose(m, 1) = choose(m, $m-1$) = m; also, in this document, choose(n, k) is 0 when k is less than 0 or greater than n.$n!$ = 1*2*3*...*n is also known as n factorial; in this document, (0!) = 1.

5 Approximations by Polynomials

This section first shows how to approximate a function on the closed unit interval, then shows how to approximate a function on *any* closed interval.

5.1 Approximations on the Closed Unit Interval

Suppose $f(\lambda)$ is continuous and maps the closed unit interval to the closed unit interval.

Then, a polynomial of a high enough degree (called n) can be used to approximate $f(\lambda)$ with an error no more than ϵ, as long as the polynomial's Bernstein coefficients can be calculated and an explicit upper bound on the approximation error is available. See my **question on MathOverflow**[7]. Examples of these polynomials (all of degree n) are given in the following table.

Name	Polynomial	Its Bernstein coefficients are found as follows:	Notes
Bernstein polynomial.	$B_n(f)$.	$f(j/n)$, where $0 \le j \le n$.	Originated with S.N. Bernstein (1912). Evaluates f at $n+1$ evenly spaced points.
Order-2 iterated Boolean sum.	$U_{n,2} = B_n(W_{n,2})$.	$W_{n,2}(j/n)$, where $0 \le j \le n$ and $W_{n,2}(\lambda) = 2f(\lambda) - B_n(f)(\lambda)$.	Micchelli (1973)[8], Guan (2009)[9], Güntürk and Li (2021, sec. 3.3)[10]. Evaluates f at $n+1$ evenly spaced points.
Order-3 iterated Boolean sum.	$U_{n,3} = B_n(W_{n,3})$.	$W_{n,3}(j/n)$, where $0 \le j \le n$ and $W_{n,3}(\lambda) = B_n(B_n(f)(\lambda)) + 3(f(\lambda) - B_n(f)(\lambda))$.	Same.

[7]https://mathoverflow.net/questions/442057/explicit-and-fast-error-bounds-for-approximating-continuous-functions

[8]Micchelli, Charles. "**The saturation class and iterates of the Bernstein polynomials**", Journal of Approximation Theory 8, no. 1 (1973): 1-18. https://www.sciencedirect.com/science/article/pii/0021904573900282

[9]Guan, Zhong. "**Iterated Bernstein polynomial approximations**", arXiv:0909.0684 (2009). https://arxiv.org/abs/0909.0684

[10]Güntürk, C.S., Li, W., "**Approximation of functions with one-bit neural networks**", arXiv:2112.09181 [cs.LG], 2021. https://arxiv.org/abs/2112.09181

Name	Polynomial	Its Bernstein coefficients are found as follows:	Notes
Butzer's linear combination (order 2).	$L_{2,n/2} = 2B_n(f(\lambda)) - B_{n/2}(f(\lambda))$.	(First, define the following operation: **Get coefficients for n given m**: Treat the coefficients $[f(0/m), f(1/m), …, f(m/m)]$ as representing a polynomial in Bernstein form of degree m, then rewrite that polynomial to one of degree n with $n+1$ Bernstein coefficients (see "**Computational Issues**"), then return those coefficients.) **Get coefficients for n given $n/2$**, call them $a[0]$, …, $a[n]$, then set the final Bernstein coefficients to $2f(j/n) - a[j]$ for each j.	Tachev (2022)[11], Butzer (1955)[12]. $n \geq 6$ must be even. Evaluates f at $n/2 + 1$ evenly spaced points.
Butzer's linear combination (order 3).	$L_{3,n/4} = B_{n/4}(f)/3 + B_n(f) \cdot 8/3 - 2B_{n/2}(f)$	**Get coefficients for n given $n/4$**, call them $a[0]$, …, $a[n]$, then **get coefficients for n given $n/2$**, call them $b[0]$, …, $b[n]$, then set the final Bernstein coefficients to $a[j]/3 - 2b[j] + 8f(j/n)/3$ for each j.	Butzer (1955)[13]. $n \geq 4$ must be divisible by 4. Evaluates f at $n/2 + 1$ evenly spaced points.

[11]Tachev, Gancho. "**Linear combinations of two Bernstein polynomials**", *Mathematical Foundations of Computing*, 2022. https://doi.org/10.3934/mfc.2022061

[12]Butzer, P.L., "Linear combinations of Bernstein polynomials", Canadian Journal of Mathematics 15 (1953).

[13]Butzer, P.L., "Linear combinations of Bernstein polynomials", Canadian Journal of Mathematics 15 (1953).

Name	Polynomial	Its Bernstein coefficients are found as follows:	Notes
Lorentz operator (order 2).	$Q_{n-2,2} = B_{n-2}(f) - x(1-x) \cdot B_{n-2}(f^{(2)})/(2(n-2))$.	**Get coefficients for n given $n-2$, call them** $a[0]$, ..., $a[n]$. Then for each integer j with $1 \le j < n$, subtract z from $a[j]$, where $z = (((f^{(2)}((j-1)/(n-2))) / (4(n-2))) \cdot 2j(n-j)/((n-1) \cdot (n))$. The final Bernstein coefficients are now $a[0]$, ..., $a[n]$.	Holtz et al. (2011)[14]; Bernstein (1932)[15]; Lorentz (1966)[16]. $n \ge 4$; $f^{(2)}$ is the second derivative of f. Evaluates f and $f^{(2)}$ at $n-1$ evenly spaced points.

The goal is now to find a polynomial of degree n, written in Bernstein form, such that—

1. the polynomial is within ϵ of $f(\lambda)$, and
2. each of the polynomial's Bernstein coefficients is not less than 0 or greater than 1 (assuming none of f's values is less than 0 or greater than 1).

For some of the polynomials given earlier, a degree n can be found so that the degree-n polynomial is within ϵ of f, if f is continuous and meets other conditions. In general, to find the degree n, solve the error bound's equation for n and round the solution up to the nearest integer. See the following table, where:

- M_r is not less than the maximum of the absolute value of f's r-th derivative.
- H_r is not less than f's r-th derivative's Hölder constant (for the specified Hölder exponent α).
- L_r is not less than f's r-th derivative's Lipschitz constant.

If $f(\lambda)$:	Then the following polynomial:	Is close to f with the following error bound:	And a value of n that achieves the bound is:	Notes
Has Hölder continuous second derivative (see "Definitions").	$U_{n,2}(f)$.	$\varepsilon = (5H_2 + 4M_2) / (32n^{1+\alpha/2})$.	$n=\max(3, \text{ceil}(((5H_2 + 4M_2) / (32\epsilon))^{2/(2+\alpha)}))$.	$n \ge 3$. $0 < \alpha \le 1$ is second derivative's Hölder exponent. See Proposition B10C in appendix.
Has Lipschitz continuous second derivative.	$U_{n,2}(f)$.	$\varepsilon = (5L_2 + 4M_2) / (32n^{3/2})$.	$n=\max(3, \text{ceil}(((5L_2 + 4M_2) / (32\epsilon))^{2/3}))$.	$n \ge 3$. Special case of previous entry.

[14]Holtz, O., Nazarov, F., Peres, Y., "**New Coins from Old, Smoothly**", *Constructive Approximation* 33 (2011). https://link.springer.com/content/pdf/10.1007/s00365-010-9108-5.pdf

[15]Bernstein, S. N. (1932). "Complément a l'article de E. Voronovskaya." CR Acad. URSS, 86-92.

[16]G.G. Lorentz, "The degree of approximation by polynomials with positive coefficients", 1966.

If $f(\lambda)$:	Then the following polynomial:	Is close to f with the following error bound:	And a value of n that achieves the bound is:	Notes
Has continuous third derivative.	$L_{2,n/2}(f)$.	$\varepsilon = (3*\text{sqrt}(3-4/n)/4)*M_3/n^2 < (3*\text{sqrt}(3)/4)*M_3/n^2 < 1.29904*M_3/n^2 \le 1.29904*M_3/n^{3/2}$.	n=max(6,ceil($\frac{3^{3/4}\sqrt{M_3/\epsilon}}{2}$)) ≤ max(6,ceil((113976/100000) * sqrt(M_3/ε))) ≤ max(6, ceil((($1.29904 M_3$)/ϵ)$^{2/3}$)). (If n is now odd, add 1.)	Tachev (2022)[17]. $n \ge 6$ must be
Has Hölder continuous third derivative.	$U_{n,2}(f)$.	$\varepsilon = (9H_3+8M_2+8M_3)/(64n^{(3+\alpha)/2})$.	n=max(6, ceil((($9H_3+8M_2+8M_3$)/(64ϵ))$^{2/(3+\alpha)}$)).	$n \ge 6$. $0 < \alpha \le 1$ is third derivative's Hölder exponent. See Proposition B10D in appendix.
Has Lipschitz continuous third derivative.	$U_{n,2}(f)$.	$\varepsilon = (9H_3+8M_2+8M_3)/(64n^2)$.	n=max(6, ceil((($9H_3+8M_2+8M_3$)/(64ϵ))$^{1/2}$)).	$n \ge 6$. Special case of previous entry.
Has Lipschitz continuous third derivative.	$L_{3,n/4}(f)$.	$\varepsilon = L_3/(8*n^2)$.	n=max(4,ceil((sqrt(439)/25) * sqrt(L_3/ε))) ≤ max(4,ceil((83810/100000) * sqrt(L_3/ε))). (Round n up to nearest multiple of 4.)	$n \ge 4$ must be divisible by 4. See Proposition B10 in appendix.
Has Lipschitz-continuous derivative.	$B_n(f)$.	$\varepsilon = L_1/(8*n)$.	$n = \text{ceil}(L_1/(8*\varepsilon))$.	Lorentz (1963)[18].[19]
Has Hölder-continuous derivative.	$B_n(f)$.	$\varepsilon = H_1/(4*n^{(1+\alpha)/2})$.	$n = \text{ceil}((H_1/(4*\varepsilon))^{2/(1+\alpha)})$.	Schurer and Steutel (1975)[20]. $0 < \alpha \le 1$ is derivative's Hölder exponent.
Is Hölder continuous.	$B_n(f)$.	$\varepsilon = H_0*(1/(4*n))^{\alpha/2}$.	$n = \text{ceil}((H_0/\varepsilon)^{2/\alpha}/4)$.	Kac (1938)[21]. $0 < \alpha \le 1$ is f's Hölder exponent.
Is Lipschitz continuous.	$B_n(f)$.	$\varepsilon = L_0*\text{sqrt}(1/(4*n))$.	$n = \text{ceil}((L_0)^2/(4*\varepsilon^2))$.	Special case of previous entry.
Is Lipschitz continuous.	$B_n(f)$.	$\varepsilon = \frac{4306+837\sqrt{6}}{5832}L_0/n^{1/2} < 1.08989L_0/n^{1/2}$.	n=ceil((L_0*1.08989/ε)^2).	Sikkema (1961)[22].

[17]Tachev, Gancho. "**Linear combinations of two Bernstein polynomials**", *Mathematical Foundations of Computing*, 2022. https://doi.org/10.3934/mfc.2022061

[18]G.G. Lorentz, "Inequalities and saturation classes for Bernstein polynomials", 1963.

[19]Qian et al. suggested an n which has the upper bound n=ceil(1+max($2n,n^2(2^nC)/\epsilon$)), where C is the maximum of f on its domain, but this is often much worse and works only if f is a polynomial (Qian, W., Riedel, M. D., & Rosenberg, I. (2011). Uniform approximation and Bernstein polynomials with coefficients in the unit interval. European Journal of Combinatorics, 32(3), 448-463).

[20]Schurer and Steutel, "On an inequality of Lorentz in the theory of Bernstein polynomials", 1975.

[21]Kac, M., "Une remarque sur les polynômes de M. S. Bernstein", Studia Math. 7, 1938.

[22]Sikkema, P.C., "Der Wert einiger Konstanten in der Theorie der Approximation mit Bernstein-Polynomen", 1961.

Bernstein polynomials ($B_n(f)$) have the advantages that only one Bernstein coefficient has to be found per run and that the coefficient will be bounded by 0 and 1 if $f(\lambda)$ is. But their disadvantage is that they approach f slowly in general, at a rate no faster than a rate proportional to $1/n$ (Voronovskaya 1932)[23].

On the other hand, polynomials other than Bernstein polynomials can approach f faster in many cases than Bernstein polynomials, but are not necessarily bounded by 0 and 1. For these polynomials, the following process can be used to calculate the required degree n, given an error tolerance of ϵ, assuming none of f's values is less than 0 or greater than 1.

1. Determine whether f is described in the preceding table. Let A be the minimum of f on the closed unit interval and let B be the maximum of f there.
2. If $0 < A \leq B < 1$, calculate n as given in the preceding table, but with $\epsilon = \min(\epsilon, A, 1 - B)$, and stop.
3. Propositions B1, B2, and B3 in the **appendix** give conditions on f so that $W_{n,2}$ or $W_{n,3}$ (as the case may be) will be nonnegative. If B is less than 1 and any of those conditions is met, calculate n as given in the preceding table, but with $\epsilon = \min(\epsilon, 1 - B)$. (For B3, set n to $\max(n, m)$, where m is given in that proposition.) Then stop; $W_{n,2}$ or $W_{n,3}$ will now be bounded by 0 and 1.
4. Calculate n as given in the preceding table. Then, if any Bernstein coefficient of the resulting polynomial is less than 0 or greater than 1, double the value of n until this condition is no longer true.

The resulting polynomial of degree n will be within ϵ of $f(\lambda)$.

Notes:

1. A polynomial's Bernstein coefficients can be rounded to multiples of δ (where $0 < \delta \leq 1$) by setting either—

   - c=floor(c/δ) * δ (rounding down), or
   - c=floor$(c/\delta + 1/2)$ * δ (rounding to the nearest multiple),

   for each Bernstein coefficient c. The new polynomial will differ from the old one by at most δ. (Thus, to find a polynomial with multiple-of-δ Bernstein coefficients that approximates f with error ϵ [which must be greater than δ], first find a polynomial with error $\epsilon - \delta$, then round that polynomial's Bernstein coefficients as given here.)

2. *Gevrey's hierarchy* is a class of "smooth" functions with known bounds on their derivatives. A function $f(\lambda)$ belongs in *Gevrey's hierarchy* if there are values $B \geq 1$, $l \geq 1$, $\gamma \geq 1$ such that f's n-th derivative's absolute value is not greater than $Bl^n n^{\gamma n}$ for every $n \geq 1$ (Kawamura et al. 2015[24]; see also (Gevrey 1918)[25]). In this case, for each $n \geq 1$—

   - the n-th derivative of f is continuous and has a maximum absolute value of at most $Bl^n n^{\gamma n}$, and
   - the $(n - 1)$-th derivative of f is Lipschitz continuous with Lipschitz constant at most $Bl^n n^{\gamma n}$.

   Gevrey's hierarchy with $\gamma = 1$ is the class of functions equaling power series (see note in next section).

5.2 Taylor Polynomials for "Smooth" Functions

Every continuous function defined on the closed interval $[a, b]$ can be written as—

[23]E. Voronovskaya, "Détermination de la forme asymptotique d'approximation des fonctions par les polynômes de M. Bernstein", 1932.

[24]Kawamura, Akitoshi, Norbert Müller, Carsten Rösnick, and Martin Ziegler. "**Computational benefit of smoothness: Parameterized bit-complexity of numerical operators on analytic functions and Gevrey's hierarchy.**" Journal of Complexity 31, no. 5 (2015): 689-714. https://doi.org/10.1016/j.jco.2015.05.001

[25]M. Gevrey, "Sur la nature analytique des solutions des équations aux dérivées partielles", 1918.

$$f(\lambda) = R_{f,r}(\lambda, x_0) + f(x_0) + f^{(1)}(x_0)\frac{(\lambda - x_0)^1}{1!} + \ldots + f^{(r)}(x_0)\frac{(\lambda - x_0)^r}{r!}$$

$$= R_{f,r}(\lambda, x_0) + Q_{f,r}(\lambda, x_0),$$

as long as the function's r-th derivative ($r \geq 0$) is defined at x_0, where $a \leq x_0 \leq b$. When this is the case, f equals—

- $Q_{f,r}(\lambda, x_0)$, the r-th *Taylor polynomial* centered at x_0, plus
- $R_{f,r}(\lambda, x_0)$, the r-th *Taylor remainder.*

If $f(\lambda)$ is "smooth" enough on the closed unit interval $[0, 1]$, and if ϵ is big enough, then Taylor's theorem shows there is a Taylor polynomial of f that comes within ϵ of f. In this section f may but need not be writable as a power series (see note).

In this section, M_r is not less than the maximum of the absolute value of f's r-th derivative.

Let $n \geq 0$ be an integer, and let $f^{(i)}$ be the i-th derivative of $f(\lambda)$. Suppose that—

1. f is continuous on the closed unit interval, and
2. f satisfies $\epsilon \leq f(0) \leq 1 - \epsilon$ and $\epsilon \leq f(1) \leq 1 - \epsilon$, and
3. f satisfies $\epsilon < f(\lambda) < 1 - \epsilon$ whenever $0 < \lambda < 1$, and
4. f's $(n + 1)$-th derivative is continuous and satisfies $\epsilon \geq M_{n+1}/((n + 1)!)$, and
5. $f(0)$ is known as well as $f^{(1)}(0), \ldots, f^{(n)}(0)$.

Then the n-th *Taylor polynomial* centered at 0 ($Q_{f,n}$) is within ϵ of f.

Items 2 and 3 above are not needed to find a polynomial within ϵ of f, but they *are* needed to ensure the Taylor polynomial's Bernstein coefficients will lie in the closed unit interval, as described after the note.

> **Note:** If $f(\lambda)$ can be rewritten as a *power series*, namely $f(\lambda) = c_0\lambda^0 + c_1\lambda^1 + \ldots + c_i\lambda^i + \ldots$ whenever $0 \leq \lambda \leq 1$ (so that f has a continuous k-th derivative for every k), and if the power series coefficients c_i—
>
> - are each greater than 0,
> - form a nowhere increasing sequence (example: $(1/4, 1/8, 1/8, 1/16, \ldots)$), and
> - meet the so-called "ratio test",
>
> then the algorithms in Carvalho and Moreira (2022)[26] can be used to find the first $n+1$ power series coefficients such that $P(\lambda)$ is within ϵ of f (see also the appendix).

Now, the Taylor polynomial $Q_{f,n}$, when written in its "power" form or "monomial" form, has "power" coefficients equal to $f(0), f^{(1)}(0)/(1!), \ldots, f^{(n)}(0)/(n!)$.

Now, rewrite $P(\lambda)$ as a polynomial in Bernstein form. (See "**Computational Issues**" for details.) Let $b_0, \ldots, b_n$ be the resulting Bernstein coefficients. If any of those Bernstein coefficients is less than 0 or greater than 1, then—

- double the value of n, then
- rewrite the Bernstein coefficients of degree $n/2$ to the corresponding Bernstein coefficients of degree n,

until none of the Bernstein coefficients is less than 0 or greater than 1.

The result will be a polynomial of degree n with $(n + 1)$ Bernstein coefficients.

---

[26]Tsai, Yi-Feng, and Rida T. Farouki. "Algorithm 812: BPOLY: An object-oriented library of numerical algorithms for polynomials in Bernstein form." ACM Transactions on Mathematical Software (TOMS) 27.2 (2001): 267-296.

5.3 Approximations on Any Closed Interval

Now, let $g(\lambda)$ be continuous on the closed interval $[a, b]$. This section shows how to adapt the previous two sections to approximate g on the interval, to the user-defined error tolerance ϵ, by a polynomial in Bernstein form on the interval $[a, b]$.

Any polynomial $p(\lambda)$ can be written in *Bernstein form on the interval* $[a, b]$ as—

$$p(\lambda) = \frac{1}{(b-a)^n}(\binom{n}{0}(\lambda - a)^0 (b - \lambda)^{n-0} a[0] +$$

$$\binom{n}{1}(\lambda - a)^1 (b - \lambda)^{n-1} a[1] + ... +$$

$$\binom{n}{n}(\lambda - a)^n (b - \lambda)^{n-n} a[n]),$$

where n is the polynomial's *degree* and $a[0]$, $a[1]$, ..., $a[n]$ are its n plus one *Bernstein coefficients for the interval* $[a, b]$ (Bărbosu 2020)[27].

The necessary changes are as follows:

- In the previous two sections, define f, M_r, a_i, and L_r as follows:
  - $f(\lambda) = g(a + (b-a)\lambda)$. This will make f continuous on the closed unit interval.
  - M_r is not less than $(b-a)^r$ times the maximum of the absolute value of g's r-th derivative on $[a, b]$.
  - L_r is not less than $(b-a)^{r+1}$ times the Lipschitz constant of g's r-th derivative on $[a, b]$.
  - $a_i = (b-a)^i f^{(i)}(0)/(i!)$.

(The error bounds that rely on H_r won't work for the time being unless $[a, b]$ is the closed unit interval.)

The result will be in the form of Bernstein coefficients for the interval $[a, b]$ rather than the interval $[0, 1]$.

> **Note:** The following statements can be shown. Let $g(x)$ be continuous on the interval $[a, b]$, and let $f(x) = g(a + (b-a)x)$.
>
> - If the r-th derivative of g is continuous and has a maximum absolute value of M on the interval, where $r \geq 1$, then the r-th derivative of $f(x)$ has a maximum absolute value of $M(b-a)^r$ on the interval $[0, 1]$.
> - If the r-th derivative of g is Lipschitz continuous with Lipschitz constant L on the interval, where $r \geq 0$, then the r-th derivative of $f(x)$ is Lipschitz continuous with Lipschitz constant $L(b-a)^{r+1}$ on the interval $[0, 1]$.
>
> **Example:** Suppose $g(x)$ is defined on the interval $[1, 3]$ and has a Lipschitz-continuous derivative with Lipschitz constant L. Let $f(x) = g(1 + (3 - 1)x)$. Then $f(x)$ has a Lipschitz-continuous derivative with Lipschitz constant $L(3 - 1)^{r+1} = L(3 - 1)^2 = 4L$ (where r is 1 in this case). Further, the Bernstein polynomial $B_n(f)$ admits the following error bound ϵ and a degree n that achieves the error tolerance ϵ: $\epsilon = (4L) \cdot 1/(8n)$ and $n = \text{ceil}((4L) \cdot 1/(8\epsilon))$. (Compare with the row starting with "Has Lipschitz-continuous derivative" in the previous section.) The error bound carries over to $g(x)$ on the interval $[1, 3]$.

---

[27]Tsai, Yi-Feng, and Rida T. Farouki. "Algorithm 812: BPOLY: An object-oriented library of numerical algorithms for polynomials in Bernstein form." ACM Transactions on Mathematical Software (TOMS) 27.2 (2001): 267-296.

5.4 Approximating an Integral

Roughly speaking, the *integral* of $f(x)$ on an interval $[a, b]$ is the "area under the graph" of that function when the function is restricted to that interval. If f is continuous there, this is the value that $\frac{1}{n}(f(a + (b - a)(1 - \frac{1}{2})/n) + f(a + (b - a)(2 - \frac{1}{2})/n) + ... + f(a + (b - a)(n - \frac{1}{2})/n))$ approaches as n gets larger and larger.

If a polynomial is in Bernstein form of degree n, and is defined on the closed unit interval:

- The polynomial's integral on the closed unit interval is equal to the average of its $(n + 1)$ Bernstein coefficients; that is, the integral is found by adding those coefficients together, then dividing by $(n + 1)$ (Tsai and Farouki 2001, section 3.4)[28].[29]

If a polynomial is in Bernstein form on the interval $[a, b]$, of degree n:

- The polynomial's integral on $[a, b]$ is found by adding the polynomial's Bernstein coefficients for $[a, b]$ together, then multiplying by $(b - a)/(n + 1)$.

Let $P(\lambda)$ be a continuous function (such as a polynomial) on the interval $[a, b]$, and let $f(\lambda)$ be a function made up of multiple continuous functions defined on a finite number of "pieces", or nonempty subintervals, that together make up the interval $[a, b]$.

- If P is within ϵ of f at every point on the interval, then its integral is within $\epsilon \times (b - a)$ of f's integral on that interval.
- If P is within $\epsilon/(b - a)$ of f at every point on the interval, then its integral is within ϵ of f's integral on that interval.

  Note: A pair of articles by Konečný and Neumann discuss approximating the integral (and maximum) of a class of functions efficiently using polynomials or piecewise functions with polynomials as the pieces: Konečný and Neumann (2021)[30]; Konečný and Neumann (2019)[31].

  Muñoz and Narkawicz (2013)[32] also discuss finding the minimum and maximum of a polynomial in Bernstein form — indeed, a polynomial is bounded above by its highest Bernstein coefficient and below by its lowest.

5.5 Approximating a Derivative

For the time being, this section works only if $f(\lambda)$ is defined on the closed unit interval, rather than an arbitrary closed interval.

If $f(\lambda)$ has a continuous r-th derivative on the closed unit interval (where r is 1 or greater), it's possible to approximate f's r-th derivative as follows:

1. Build a polynomial in Bernstein form of a degree n that is high enough such that the r-th derivative is close to f's r-th derivative with an error no more than ϵ (where ϵ is the user-defined error tolerance. See the following table.

2. Let $a[0], ..., a[n]$ be the polynomial's Bernstein coefficients. Now, to compute the polynomial's r-th derivative, do the following r times or until the process stops, whichever happens first (Tsai and

---

[28]Tsai, Yi-Feng, and Rida T. Farouki. "Algorithm 812: BPOLY: An object-oriented library of numerical algorithms for polynomials in Bernstein form." ACM Transactions on Mathematical Software (TOMS) 27.2 (2001): 267-296.

[29]Tsai, Yi-Feng, and Rida T. Farouki. "Algorithm 812: BPOLY: An object-oriented library of numerical algorithms for polynomials in Bernstein form." ACM Transactions on Mathematical Software (TOMS) 27.2 (2001): 267-296.

[30]Konečný, Michal, and Eike Neumann. "Representations and evaluation strategies for feasibly approximable functions." Computability 10, no. 1 (2021): 63-89. Also in arXiv: **1710.03702**. https://arxiv.org/abs/1710.03702

[31]Konečný, Michal, and Eike Neumann. "**Implementing evaluation strategies for continuous real functions**", arXiv:1910.04891 (2019). https://arxiv.org/abs/1910.04891

[32]Muñoz, César, and Anthony Narkawicz. "Formalization of Bernstein polynomials and applications to global optimization." Journal of Automated Reasoning 51, no. 2 (2013): 151-196.

Farouki 2001, section 3.4)[33].

- If n is 0, set $a[0] = 0$ and stop.
- For each integer k with $0 \le k \le n - 1$, set $a[k] = n \cdot (a[k+1] - a[k])$.
- Set n to $n - 1$.

3. The result is a degree-n polynomial, with Bernstein coefficients $a[0], ..., a[n]$, that approximates the r-th derivative of $f(\lambda)$.

In the following table:

- M_r is not less than the maximum of the absolute value of f's r-th derivative.
- H_r is not less than f's r-th derivative's Hölder constant (for the specified Hölder exponent α).
- L_r is not less than f's r-th derivative's Lipschitz constant.

If $f(\lambda)$:	Then the following polynomial:	Has an r-th derivative that is close to f with the following error bound:	And a value of n that achieves the bound is:	Notes
Has Hölder continuous r-th derivative.	$B_n(f)$.	$\epsilon = rM_r(r-1)/(2n) + 5H_r/(4n^{\alpha/2}) \le (rM_r(r-1)/2 + 5H_r/4)/n^{\alpha/2}$.	$n = \mathrm{ceil}(\max(r + 1, \left(\frac{(5H_r + 2M_r(r^2 - r))^2}{16\epsilon^2} \right)^{1/\alpha}))$.	Knoop and Pottinger (1976)[34]. $0 < \alpha \le 1$ is r-th derivative's Hölder exponent.

Note: In general, it is not possible to approximate a continuous function's derivative unless upper and lower bounds on the derivative are known (Konečný and Neumann (2019)[35]).

5.6 Computational Issues

Some methods in this document require rewriting a polynomial in Bernstein form of degree m to one of a higher degree n. This is also known as *degree elevation*. This rewriting works for polynomials in Bernstein form on any closed interval.

- This rewriting can be done directly in the Bernstein form, as described in Tsai and Farouki (2001, section 3.2)[36].
- This rewriting can also be done through an intermediate form called the *scaled Bernstein form* (Farouki and Rajan 1988)[37], as described in Sánchez-Reyes (2003)[38]. (A polynomial in scaled Bernstein form is also known as a *homogeneous polynomial.*)
  - The i-th Bernstein coefficient of degree m is turned to a scaled Bernstein coefficient by multiplying it by choose(m,i).
  - The i-th scaled Bernstein coefficient of degree m is turned to a Bernstein coefficient by dividing it by choose(m,i).

[33]Tsai, Yi-Feng, and Rida T. Farouki. "Algorithm 812: BPOLY: An object-oriented library of numerical algorithms for polynomials in Bernstein form." ACM Transactions on Mathematical Software (TOMS) 27.2 (2001): 267-296.

[34]Knoop, H-B., Pottinger, P., "Ein Satz vom Korovkin-Typ für C^k-Räume", Math. Zeitschrift 148 (1976).

[35]Konečný, Michal, and Eike Neumann. "**Implementing evaluation strategies for continuous real functions**", arXiv:1910.04891 (2019). https://arxiv.org/abs/1910.04891

[36]Tsai, Yi-Feng, and Rida T. Farouki. "Algorithm 812: BPOLY: An object-oriented library of numerical algorithms for polynomials in Bernstein form." ACM Transactions on Mathematical Software (TOMS) 27.2 (2001): 267-296.

[37]Farouki, Rida T., and V. T. Rajan. "**Algorithms for polynomials in Bernstein form**". Computer Aided Geometric Design 5, no. 1 (1988): 1-26. https://www.sciencedirect.com/science/article/pii/0167839688900167

[38]Sánchez-Reyes, J. (2003). **Algebraic manipulation in the Bernstein form made simple via convolutions**. Computer-Aided Design, 35(10), 959-967. https://www.sciencedirect.com/science/article/pii/S0010448503000216

Some methods in this document require rewriting a polynomial in "power" form of degree m (also known as "monomial" form) to Bernstein form of degree m. This rewriting works only for polynomials in Bernstein form on the closed unit interval.

- This rewriting can be done directly using the so-called "matrix method" from Ray and Nataraj (2012)[39].
- This rewriting can also be done by rewriting the polynomial from "power" form to scaled Bernstein form (see Sánchez-Reyes (2003, section 2.6)[40]), then converting the scaled Bernstein form to Bernstein form.

6 Approximations by Rational Functions

Consider the class of rational functions $p(\lambda)/q(\lambda)$ that map the closed unit interval to itself, where $q(\lambda)$ is in Bernstein form with nonnegative coefficients. Then rational functions of this kind are not much better than polynomials in approximating $f(\lambda)$ when—

- f has only a finite number of continuous derivatives on the open interval (0, 1) (Borwein 1979, section 3)[41], *or*
- $f(\lambda)$ is writable as $a_0\lambda^0 + a_1\lambda^1 + ...$, where $a_k \geq (k+1)a_{k+1} \geq 0$ whenever $k \geq 0$ (Borwein 1980)[42].

In addition, rational functions are not much better than polynomials in approximating $f(\lambda)$ when—

- f has only a finite number of continuous derivatives on the half-open interval (0, 1], *and*
- the rational function's denominator has no root that is a complex number whose real part is between 0 and 1 (Borwein 1979, theorem 29)[43].

Worth discussing are the approximating rational functions studied in Zhang and Liu (2022)[44] and Themistoclakis and Van Barel (2024)[45]. In the latter paper, though, it might be a bit difficult to glean error estimates of the kind given in the second table in the section "Approximations on the Closed Unit Interval", earlier in the present article. I seek help on that.

7 Request for Additional Methods

Readers are requested to let me know of additional solutions to the following problems:

1. Let $f(\lambda)$ be continuous and map the closed unit interval to itself. Given $\epsilon > 0$, and given that $f(\lambda)$ belongs to a large class of functions (for example, it has a continuous, Lipschitz continuous, concave, or nowhere decreasing k-th derivative for some integer k, or any combination of these), compute the Bernstein coefficients of a polynomial or rational function (of some degree n) that is within ϵ of $f(\lambda)$.

   The approximation error must be no more than a constant times $1/n^{r/2}$ if the specified class has only functions with continuous r-th derivative.

   Methods that use only integer arithmetic and addition and multiplication of rational numbers are preferred (thus, Chebyshev interpolants and other methods that involve cosines, sines, π, exp, and ln are not preferred).

[39]S. Ray, P.S.V. Nataraj, "**A Matrix Method for Efficient Computation of Bernstein Coefficients**", Reliable Computing 17(1), 2012. https://interval.louisiana.edu/reliable-computing-journal/volume-17/reliable-computing-17-pp-40-71.pdf

[40]Sánchez-Reyes, J. (2003). **Algebraic manipulation in the Bernstein form made simple via convolutions**. Computer-Aided Design, 35(10), 959-967. https://www.sciencedirect.com/science/article/pii/S0010448503000216

[41]Borwein, P. B. (1979). Restricted uniform rational approximations (Doctoral dissertation, University of British Columbia).

[42]Borwein, Peter B. "Approximations by rational functions with positive coefficients." Journal of Mathematical Analysis and Applications 74, no. 1 (1980): 144-151.

[43]Borwein, P. B. (1979). Restricted uniform rational approximations (Doctoral dissertation, University of British Columbia).

[44]Zhang, Ren-Jiang, and Xing Liu. "Rational interpolation operator with finite Lebesgue constant." Calcolo 59.1 (2022): 10.

[45]Themistoclakis, W., Van Barel, M. A note on generalized Floater–Hormann interpolation at arbitrary distributions of nodes. Numer Algor (2024). **https://doi.org/10.1007/s11075-024-01933-6** .

2. Find a polynomial P in Bernstein form that approximates a strictly increasing polynomial Q on the closed unit interval such that the *inverse* of P is within ϵ of the inverse of Q.

   - There is an algorithm in Farouki (2000)[46], but the algorithm is not accessible free.

3. Find a polynomial P in Bernstein form that approximates a strictly increasing real analytic function f on the closed unit interval such that the *inverse* of P is within ϵ of the inverse of f.

   (Note: There is no bounded convergence rate for P if f is assumed only to have a continuous k-th derivative for every k; a counterexample is $h(x) = \exp(-1/x)$ $(h(0) = 0)$, $h(h(x))$, $h(h(h(x)))$, and so on.)

See also the **open questions**[47].

8 References on Polynomial Sequences with "Fast" Approximation

The following references discuss schemes that—

- approximate functions with a continuous r-th derivative on the closed unit interval, where $r \geq 3$,
- using polynomials of degree n,
- at a rate no slower than a constant times $1/n^{r/2}$, and
- without introducing transcendental or trigonometric functions.

Holtz et al. (2011)[48]; Sevy (1991)[49] and references there; Waldron (2009)[50]; Costabile et al. (2005)[51]; Han (2003)[52]. Excluded from this list are schemes that employ *splines* (piecewise polynomials), or sequences of nonpolynomial functions.

There may be other useful schemes for polynomials not mentioned in this document or in the references just given. There may also be schemes that do not converge to the target function but can be made to achieve an approximation error of ϵ or less (where ϵ is a user-defined positive value).

9 Notes

10 Appendix

10.1 Results Used in Approximations by Polynomials

Lemma A1: Let—

$$f(x) = a_0 x^0 + a_1 x^1 + ...,$$

[46]Farouki, Rida T. "Convergent inversion approximations for polynomials in Bernstein form." Computer Aided Geometric Design 17.2 (2000): 179-196.

[47]https://peteroupc.github.io/bernreq.html#Polynomials_that_approach_a_factory_function_fast

[48]Holtz, O., Nazarov, F., Peres, Y., "**New Coins from Old, Smoothly**", Constructive Approximation 33 (2011). https://link.springer.com/article/10.1007/s00365-010-9108-5

[49]Sevy, J., "Acceleration of convergence of sequences of simultaneous approximants", dissertation, Drexel University, 1991.

[50]Waldron, S., "**Increasing the polynomial reproduction of a quasi-interpolation operator**", Journal of Approximation Theory 161 (2009). https://www.sciencedirect.com/science/article/pii/S0021904508001640

[51]Costabile, F., Gualtieri, M.I., Serra, S., "Asymptotic expansion and extrapolation for Bernstein polynomials with applications", *BIT* 36 (1996).

[52]Han, Xuli. "Multi-node higher order expansions of a function." Journal of Approximation Theory 124.2 (2003): 242-253. **https://doi.org/10.1016/j.jat.2003.08.001**

where the a_i are constants each 0 or greater and sum to a finite value and where $0 \leq x \leq 1$ (the domain is the closed unit interval). Then f is convex and has a maximum at 1.

Proof: By inspection, $f(x)$ is a power series and is nonnegative wherever $x \geq 0$ (and thus wherever $0 \leq x \leq 1$). Each of its terms has a maximum at 1 since—

- for $n = 0$, $a_0 x^0 = a_0$ is a nonnegative constant (which trivially reaches its maximum at 1), and
- for each n where $a_0 = 0$, $a_0 x^n$ is the constant 0 (which trivially reaches its maximum at 1), and
- for each other n, x^n is a strictly increasing function and multiplying that by a_n (a positive constant) doesn't change whether it's strictly increasing.

Since all of these terms have a maximum at 1 on the domain, so does their sum.

The derivative of f is—

$$f'(x) = 1 \cdot a_1 x^0 + ... + i \cdot a_i x^{i-1} + ...,$$

which is still a power series with nonnegative values of a_n, so the proof so far applies to f' instead of f. By induction, the proof so far applies to all derivatives of f, including its second derivative.

Now, since the second derivative is nonnegative wherever $x \geq 0$, and thus on its domain, f is convex, which completes the proof. [ ]

Proposition A2: For a function $f(x)$ as in Lemma A1, let—

$$g_n(x) = a_0 x^0 + ... + a_n x^n,$$

and have the same domain as f. Then for every $n \geq 1$, $g_n(x)$ is within ϵ of $f(x)$, where $\epsilon = f(1) - g_n(1)$.

Proof: g_n, consisting of the first $n+1$ terms of f, is a power series with nonnegative values for $a_0, ..., a_n$, so by Lemma A1, it has a maximum at $x = 1$. The same is true for $f - g_n$, consisting of the remaining terms of f. Since the latter has a maximum at $x = 1$, the maximum error is $\epsilon = f(1) - g_n(1)$. [ ]

For a function $f(x)$ described in Lemma A1, $f(1) = a_0 1^0 + a_1 1^1 + ... = a_0 + a_1 + ...$, and f's error behavior is described at $x = 1$, so the algorithms given in Carvalho and Moreira (2022)[53] — which apply to infinite sums — can be used to "cut off" f at a certain number of terms and do so with a controlled error.

Proposition B1: Let $f(\lambda)$ map the closed unit interval to itself and be continuous and concave. Then $W_{n,2}$ and $W_{n,3}$ (as defined in "For Certain Functions") are nonnegative on the closed unit interval.

Proof: For $W_{n,2}$ it's enough to prove that $B_n(f) \leq f$ for every $n \geq 1$. This is the case because of Jensen's inequality and because f is concave.

For $W_{n,3}$ it must also be shown that $B_n(B_n(f)(\lambda))$ is nonnegative. For this, using only the fact that f maps the closed unit interval to itself, $B_n(f)$ will have Bernstein coefficients in that interval (each of those coefficients is a value of f) and so will likewise map the closed unit interval to itself (Qian et al. 2011)[54]. Thus, by induction, $B_n(B_n(f)(\lambda))$ is nonnegative. The discussion for $W_{n,2}$ also shows that $(f - B_n(f))$ is nonnegative as well. Thus, $W_{n,3}$ is nonnegative on the closed unit interval. [ ]

Proposition B2: Let $f(\lambda)$ map the closed unit interval to itself, be continuous, nowhere decreasing, and subadditive, and equal 0 at 0. Then $W_{n,2}$ is nonnegative on the closed unit interval.

[53]Tsai, Yi-Feng, and Rida T. Farouki. "Algorithm 812: BPOLY: An object-oriented library of numerical algorithms for polynomials in Bernstein form." ACM Transactions on Mathematical Software (TOMS) 27.2 (2001): 267-296.

[54]Qian, Weikang, Marc D. Riedel, and Ivo Rosenberg. "Uniform approximation and Bernstein polynomials with coefficients in the unit interval." European Journal of Combinatorics 32, no. 3 (2011): 448-463.

Proof: The assumptions on f imply that $B_n(f) \le 2f$ (Li 2000)[55], showing that $W_{n,2}$ is nonnegative on the closed unit interval. [ ]

> **Note:** A subadditive function f has the property that $f(a+b) \le f(a) + f(b)$ whenever a, b, and $a+b$ are in f's domain.

Proposition B3: Let $f(\lambda)$ map the closed unit interval to itself and have a Lipschitz-continuous derivative with Lipschitz constant L. If $f(\lambda) \ge \frac{L\lambda(1-\lambda)}{2m}$ on f's domain, for some $m \ge 1$, then $W_{n,2}$ is nonnegative there, for every $n \ge m$.

Proof: Let $E(\lambda, n) = \frac{L\lambda(1-\lambda)}{2n}$. Lorentz (1963)[56] showed that with this Lipschitz derivative assumption on f, B_n differs from $f(\lambda)$ by no more than $E(\lambda, n)$ for every $n \ge 1$ and wherever $0 < \lambda < 1$. As is well known, $B_n(0) = f(0)$ and $B_n(1) = f(1)$. By inspection, $E(\lambda, n)$ is biggest when $n = 1$ and decreases as n increases. Assuming the worst case that $B_n(\lambda) = f(\lambda) + E(\lambda, m)$, it follows that $W_{n,2} = 2f(\lambda) - B_n(\lambda) \ge 2f(\lambda) - f(\lambda) - E(\lambda, m) = f(\lambda) - E(\lambda, m) \ge 0$ whenever $f(\lambda) \ge E(\lambda, m)$. Because $E(\lambda, k+1) \le E(\lambda, k)$ for every $k \ge 1$, the preceding sentence holds true for every $n \ge m$. [ ]

The following results deal with useful quantities when discussing the error in approximating a function by Bernstein polynomials. Suppose a coin shows heads with probability p, and n independent tosses of the coin are made, where n is 1 or greater. Then the total number of heads X follows a *binomial distribution*, and the r-th central moment of that distribution is as follows:

$$T_{n,r}(p) = \mathbb{E}[(X - \mathbb{E}[X])^r] = \sum_{k=0}^{n}(k - np)^r \binom{n}{k}p^k(1-p)^{n-k},$$

where $\mathbb{E}[.]$ is the expected value ("long-run average").

- Traditionally, the central moment of X/n or the ratio of heads to tosses is denoted $S_{n,r}(p) = T_{n,r}(p)/n^r = \mathbb{E}[(X/n - \mathbb{E}[X/n])^r]$. ($T$ and S are notations of S.N. Bernstein, known for Bernstein polynomials.)
- The r-th *central absolute moment* of X/n or the ratio of heads to tosses is denoted $M_{n,r}(p) = \mathbb{E}[\mathrm{abs}(X/n - \mathbb{E}[X/n])^r] = B_n(\mathrm{abs}(\lambda - p)^r)(p)$. If r is even, $M_{n,r}(p) = S_{n,r}(p)$.

The following results bound the absolute value of $T_{n,r}$, $S_{n,r}$, and $M_{n,r}$.[57]

Lemma B5: For every integer $n \ge 1$:

- $\mathrm{abs}(S_{n,0}(p)) = 1 = 1 \cdot (p(1-p)/n)^{0/2}$.
- $\mathrm{abs}(S_{n,1}(p)) = 0 = 0 \cdot (p(1-p)/n)^{1/2}$.
- $\mathrm{abs}(S_{n,2}(p)) = p(1-p)/n = 1 \cdot (p(1-p)/n)^{2/2}$.

The proof is straightforward.

Result B5A: Let $\Delta_n(x) = \max(1/n, (x(1-x)/n)^{1/2})$. For every real number $r > 0$, $M_{n,r}(p) \le (c + d)(\Delta_n(x))^r$, where $c = 2 \cdot 4^{r/2}\Gamma(r/2 + 1)$, $d = 2 \cdot 8^r\Gamma(r+1)$, and $\Gamma(x)$ is the gamma function.

Proof: By Theorem 1 of Adell et al. (2015)[58] with $\delta = 1/2$, $M_{n,r}(p) \le c(p(1-p)/n)^{r/2} + d/n^r$, and in turn, $c(p(1-p)/n)^{r/2} + d/n^r \le c(\Delta_n(p))^r + d(\Delta_n(p))^r = (c+d)(\Delta_n(p))^r$. [ ]

By Result B5A, $c + d = 264$ when $r = 2$, $c + d < 6165.27$ when $r = 3$, and $c + d = 196672$ when $r = 4$.

---

[55]Li, Zhongkai. "Bernstein polynomials and modulus of continuity." Journal of Approximation Theory 102, no. 1 (2000): 171-174.

[56]G.G. Lorentz, "Inequalities and saturation classes for Bernstein polynomials", 1963.

[57]*Summation notation*, involving the Greek capital sigma (Σ), is a way to write the sum of one or more terms of similar form. For example, $\sum_{k=0}^{n} g(k)$ means $g(0) + g(1) + ... + g(n)$, and $\sum_{k\ge 0} g(k)$ means $g(0) + g(1) + ...$

[58]Adell, J. A., Bustamante, J., & Quesada, J. M. (2015). Estimates for the moments of Bernstein polynomials. Journal of Mathematical Analysis and Applications, 432(1), 114-128.

Lemma B7: Let $r \geq 0$, and let $\sigma(r, t) = (r!)/(((r/2)!)t^{r/2})$. Then for values of r and n described in the following table, $M_{n,r}(p) \leq \mu_{n,r}/n^{r/2}$, where $\mu_{n,r}$ is as given in the table.

If r...	Then $\mu_{n,r}$ is...
Is an even integer.	$\sigma(r, 6)$, for every integer $n \geq 1$.
Is an even integer, but not greater than 44.	$\sigma(r, 8)$, for every integer $n \geq 1$.
Is 1.	$1/2$, for every integer $n \geq 1$.
Is odd, and $3 \leq r \leq 43$.	$\sqrt{\sigma(r-1, 8)\sigma(r+1, 8)} = r^{1/2}(r-1)!/(2 \cdot 8^{(r-1)/2}((r-1)/2)!)$, for every integer $n \geq 2$.
Is odd and greater than 43.	$\sqrt{\sigma(r-1, 6)\sigma(r+1, 6)}$, for every integer $n \geq 2$.

Proof: The first row comes from a result of Adell and Cárdenas-Morales (2018)[59]. The second row is an improved result of the first, from Molteni (2022)[60]. The third row follows from Lemma B5 and Schwarz's inequality; for $n = 1$, also follows from Cheng (1983)[61]. The fourth and fifth rows follow from the first and second as well as that the absolute central moment for odd r can be bounded for every integer $n \geq 2$, using **Schwarz's inequality**[62] (see also Bojanić and Shisha 1975[63] for the case $r = 4$). [ ]

Taylor polynomials and Taylor remainders were discussed in the section "Taylor Polynomials for 'Smooth' Functions". The following lemma gives bounds on the Taylor remainder's Bernstein polynomial.

Lemma B9: Let $r \geq 0$ be an integer, and let x_0 satisfy $0 \leq x_0 \leq 1$. Let $f(\lambda)$ have a Lipschitz-continuous r-th derivative on the closed unit interval (see "**Definitions**"), with Lipschitz constant M or less. Denote $B_n(f)$ as the Bernstein polynomial of f of degree n. Then the following bound holds true: $\mathrm{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \leq (M\mu_{r+1})/(((r+1)!)n^{(r+1)/2})$ for every integer $n \geq 2$ (and also for $n = 1$ if r is odd), where μ_r is as defined in Lemma B7.

Proof: This result relies on Lemma 2C in the article "**Supplemental Notes for Bernoulli Factory Algorithms**[64]", with $Y = X/n$, where X is a binomial random variable with n tosses and heads probability x_0; Y takes only values on the closed unit interval. [ ]

> **Note:** It would be interesting to strengthen this lemma, at least for $r \leq 10$, with a bound of the form $MC \cdot \max(1/n, (x_0(1-x_0)/n)^{1/2})^{r+1}$, where C is an explicitly given constant depending on r, which is possible because the Bernstein polynomial of $\mathrm{abs}(\lambda - x_0)^{r+1}$ can be bounded in this way (Lorentz 1966)[65].

Corollary B9A: Let $f(\lambda)$ have a Lipschitz-continuous r-th derivative on the closed unit interval, and let M be that r-th derivative's Lipschitz constant or greater. Let $R_{f,r}(\lambda, x_0)$ be as in Lemma B9. Then, for every $0 \leq x_0 \leq 1$:

If r is:	Then $\mathrm{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \leq$ ...
0.	$M(1/2)/n^{1/2}$ for every integer $n \geq 1$.
1.	$M(1/8)/n = 0.125M/n$ for every integer $n \geq 1$.
2.	$M(\sqrt{3}/48)/n^{3/2} < 0.3609M/n^{3/2}$ for every integer $n \geq 2$.

[59]Adell, J.A., Cárdenas-Morales, D., "**Quantitative generalized Voronovskaja's formulae for Bernstein polynomials**", Journal of Approximation Theory 231, July 2018. https://www.sciencedirect.com/science/article/pii/S0021904518300376

[60]Molteni, Giuseppe. "Explicit bounds for even moments of Bernstein's polynomials." Journal of Approximation Theory 273 (2022): 105658.

[61]Cheng, F., "On the rate of convergence of Bernstein polynomials of functions of bounded variation", Journal of Approximation Theory 39 (1983).

[62]https://mathworld.wolfram.com/SchwarzsInequality.html

[63]G.G. Lorentz, *Bernstein polynomials*, 1953.

[64]https://peteroupc.github.io/bernsupp.html

[65]G.G. Lorentz, "The degree of approximation by polynomials with positive coefficients", 1966.

If r is:	Then $\text{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \leq \ldots$
3.	$M(1/128)/n^2 = 0.0078125M/n^2$ for every integer $n \geq 1$.
4.	$M(\sqrt{5}/1280)/n^{5/2} < 0.001747/n^{5/2}$ for every integer $n \geq 2$.
5.	$M(1/3072)/n^3 < 0.0003256/n^3$ for every integer $n \geq 1$.

Proposition B10: Let $f(\lambda)$ have a Lipschitz-continuous third derivative on the closed unit interval. For each $n > 4$ that is divisible by 4, let $L_{3,n/4}(f) = (1/3) \cdot B_{n/4}(f) - 2 \cdot B_{n/2}(f) + (8/3) \cdot B_n(f)$. Then $L_{3,n/4}(f)$ is within $\frac{\Lambda_3}{8n^2}$ of f, where Λ_3 is the maximum of that third derivative's Lipschitz constant or greater.

Proof: This proof is inspired by the proof technique in Tachev (2022)[66].

Because f has a Lipschitz-continuous third derivative, f has the Taylor remainder $R_{f,3}(\lambda, x_0)$ given in Lemma B9 and Corollary B9A.

It is known that $L_{3,n/4}$ is a linear operator that preserves polynomials of degree 3 or less (cubic, quadratic, and linear functions, and constants), so that $L_{3,n/4}(f) = f$ whenever f is a polynomial of degree 3 or less (Ditzian and Totik 1987)[67], Butzer (1955)[68], May (1976)[69]. Because of this, it can be assumed without loss of generality that f and the first, second, and third derivatives of f equal zero at x_0.

Therefore—

$$\text{abs}(L_{3,n/4}(f(\lambda))(x_0) - f(x_0)) = \text{abs}(L_{3,n/4}(R_{f,3}(\lambda, x_0))).$$

Now denote σ_n as the maximum of $\text{abs}(B_n(R_{f,3}(\lambda, x_0))(x_0))$ over $0 \leq x_0 \leq 1$. By Corollary B9A, $\sigma_n \leq \Lambda_3(1/128)/n^{3/2}$. Therefore—

$$\text{abs}(L_{3,n/4}(R_{f,3}(\lambda, x_0))) \leq (1/3) \cdot \sigma_{n/4} + 2 \cdot \sigma_{n/2} + (8/3) \cdot \sigma_n \leq \frac{\Lambda_3}{8n^2}.$$

[ ]

The proof of Proposition B10 shows how to prove an upper bound on the approximation error for polynomials written as—

$$P(f)(x) = \alpha_0 B_{n(0)}(f)(x) + \alpha_1 B_{n(1)}(f)(x) + \ldots + \alpha_k B_{n(k)}(f)(x)$$

(where α_i are real numbers and $n(i) \geq 2$ is an integer), as long as P preserves all polynomials of degree r or less and f has a Lipschitz-continuous r-th derivative. An example is the polynomials $T_q^{(0)}$ described in Costabile et al. (1996)[70], citing Stoer and Bulirsch (1970)[71].

The following error bounds, which make use of Corollary B9A and the proof technique in Proposition B10, can be shown. In the following table, Λ_r is the maximum of f's r-th derivative's Lipschitz constant or greater, and each result applies only to values of n where all the numbers in the third column are integers greater than 1.

[66]Tachev, Gancho. "**Linear combinations of two Bernstein polynomials**", *Mathematical Foundations of Computing*, 2022. https://doi.org/10.3934/mfc.2022061

[67]Ditzian, Z., Totik, V., *Moduli of Smoothness*, 1987.

[68]Butzer, P.L., "Linear combinations of Bernstein polynomials", Canadian Journal of Mathematics 15 (1953).

[69]May, C.P., "Saturation and inverse theorems for a class of exponential-type operators", Canadian Journal of Mathematics 28 (1976).

[70]Costabile, F., Gualtieri, M.I., Serra, S., "Asymptotic expansion and extrapolation for Bernstein polynomials with applications", *BIT* 36 (1996).

[71]Stoer, J., Bulirsch, R., *Introduction to Numerical Analysis*, 1970.

Property of f on the closed unit interval	α_0, α_1, ...	$n(0)$, $n(1)$, ...	Upper bound of error
Has a Lipschitz-continuous second derivative.	-1, 2	$\frac{n}{2}$, n	$\frac{\sqrt{3}+\sqrt{6}}{24n^{3/2}} < \frac{0.1743}{n^{3/2}}$
Has a Lipschitz-continuous third derivative.	$\frac{1}{3}$, -2, $\frac{8}{3}$	$\frac{n}{4}$, $\frac{n}{2}$, n	$\frac{1}{8n^2} = \frac{0.125}{n^2}$
Has a Lipschitz-continuous fourth derivative.	$-\frac{1}{21}$, $\frac{2}{3}$, $-\frac{8}{3}$, $\frac{64}{21}$	$\frac{n}{8}$, $\frac{n}{4}$, $\frac{n}{2}$, n	$\frac{11\sqrt{10}+16\sqrt{5}}{840n^{5/2}} < \frac{0.08401}{n^{5/2}}$
Has a Lipschitz-continuous fifth derivative.	$\frac{1}{315}$, $-\frac{2}{21}$, $\frac{8}{9}$, $-\frac{64}{21}$, $\frac{1024}{315}$	$\frac{n}{16}$, $\frac{n}{8}$, $\frac{n}{4}$, $\frac{n}{2}$, n	$\frac{1}{21n^3} < \frac{0.04762}{n^3}$
Has a Lipschitz-continuous third derivative.	$\frac{1}{2}$, -4, $\frac{9}{2}$	$\frac{n}{3}$, $\frac{n}{2}$, n	$\frac{25}{128n^2} = 0.1953125/n^2$
Has a Lipschitz-continuous fourth derivative.	$-\frac{1}{6}$, 4, $-\frac{27}{2}$, $\frac{32}{3}$	$\frac{n}{4}$, $\frac{n}{3}$, $\frac{n}{2}$, n	$\frac{8\sqrt{5}+18\sqrt{15}+27\sqrt{10}}{640n^{5/2}} < \frac{0.2703}{n^{2.5}}$
Has a Lipschitz-continuous fifth derivative.	$\frac{1}{24}$, $-\frac{8}{3}$, $\frac{81}{4}$, $-\frac{128}{3}$, $\frac{625}{24}$	$\frac{n}{5}$, $\frac{n}{4}$, $\frac{n}{3}$, $\frac{n}{2}$, n	$\frac{545}{1536n^3} < \frac{0.3549}{n^3}$

The *Lorentz operator* of order 2 is denoted as $Q_{n,2}(f) = B_n(f)(x) - \frac{x(1-x)}{2n}B_n(f^{(2)})(x)$ (Holtz et al. 2011)[72], (Lorentz 1966)[73]. This operator is a polynomial in Bernstein form of degree $n+2$.

Proposition B10A: Let $f(\lambda)$ have a Lipschitz-continuous second derivative on the closed unit interval. If $n \geq 2$ is an integer, $Q_{n,2}(f)$ is within $\frac{L_2(\sqrt{3}+3)}{48n^{3/2}} < 0.098585L_2/(n^{3/2})$ of f, where L_2 is the maximum of that second derivative's Lipschitz constant or greater.

Proof: Since $Q_{n,2}(f)$ preserves polynomials of degree 2 or less (quadratic, linear, and constant functions) (Holtz et al. 2011, Lemma 14)[74] and since f has a Lipschitz-continuous second derivative, f has the Lagrange remainder $R_{f,2}(\lambda, x_0)$ given in Lemma B9, and $f^{(2)}$, the second derivative of f, has the Lagrange remainder $R_{f^{(2)},0}(\lambda, x_0)$. Thus, using Corollary B9A, the error bound can be written as—

$$\text{abs}(Q_{n,2}(f(\lambda))(x_0) - f(x_0)) \leq \text{abs}(B_n(R_{f,2}(\lambda, x_0))) + \frac{x_0(1-x_0)}{2n}\text{abs}(B_n(R_{f^{(2)},0}(\lambda, x_0)))$$

$$\leq \frac{\sqrt{3}L_2}{48n^{3/2}} + \frac{1}{8n}\frac{L_2}{2n^{1/2}} = \frac{L_2(\sqrt{3}+3)}{48n^{3/2}} < 0.098585L_2/(n^{3/2}).$$

[ ]

Corollary B10B: Let $f(\lambda)$ have a continuous second derivative on the closed unit interval. Then $B_n(f)$ is within $\frac{M_2}{8n}$ of f, where M_2 is the maximum of that second derivative's absolute value or greater.

[72]Holtz, O., Nazarov, F., Peres, Y., "**New Coins from Old, Smoothly**", *Constructive Approximation* 33 (2011). https://link.springer.com/content/pdf/10.1007/s00365-010-9108-5.pdf

[73]G.G. Lorentz, "The degree of approximation by polynomials with positive coefficients", 1966.

[74]Holtz, O., Nazarov, F., Peres, Y., "**New Coins from Old, Smoothly**", *Constructive Approximation* 33 (2011). https://link.springer.com/content/pdf/10.1007/s00365-010-9108-5.pdf

Proof: Follows from Lorentz (1963)[75] and the well-known fact that M_2 is an upper bound of f's first derivative's (minimal) Lipschitz constant. [ ]

In the following propositions:

- $f^{(r)}$ means the r-th derivative of the function f.
- $M_r = \max(\text{abs}(f^{(r)}))$ means a value equal to or greater than the maximum of the absolute value of $f^{(r)}$.
- H_r means a value equal to or greater than the Hölder constant of $f^{(r)}$.

Proposition B10C: Let $f(\lambda)$ have a Hölder-continuous second derivative on the closed unit interval, with Hölder exponent α ($0 < \alpha \leq 1$). Let $U_{n,2}(f) = B_n(2f - B_n(f))$ be f's iterated Boolean sum of order 2 of Bernstein polynomials. Then if $n \geq 3$ is an integer, the error in approximating f with $U_{n,2}(f)$ is as follows:

$$\text{abs}(f - U_{n,2}(f)) \leq \frac{M_2}{8n^2} + 5H_2/(32n^{1+\alpha/2}) \leq ((5H_2 + 4M_2)/32)/n^{1+\alpha/2}.$$

Proof: This proof is inspired by a result in Draganov (2014, Theorem 4.1)[76].

The error to be bounded can be expressed as $\text{abs}((B_n(f) - f)(B_n(f) - f))$. Following Corollary B10B:

$$\text{abs}((B_n(f) - f)(B_n(f) - f)) \leq \frac{1}{8n} \max(\text{abs}((B_n(f))^{(2)} - f^{(2)})). \tag{B10C-1}$$

It thus remains to estimate the right-hand side of the bound. A result by Knoop and Pottinger (1976)[77], which works for every $n \geq 3$, is what is known as a *simultaneous approximation* error bound, showing that the second derivative of the Bernstein polynomial approaches that of f as n increases. Using this result:

$$\text{abs}((B_n(f))^{(2)} - f^{(2)}) \leq \frac{1}{n}M_2 + (5/4)H_2/n^{\alpha/2},$$

so—

$$\text{abs}((B_n(f) - f)(B_n(f) - f)) \leq \frac{1}{8n}\left(\frac{1}{n}M_2 + (5/4)H_2/n^{\alpha/2}\right)$$

$$\leq \frac{M_2}{8n^2} + \frac{5H_2}{32n^{1+\alpha/2}} \leq \frac{5H_2 + 4M_2}{32}\frac{1}{n^{1+\alpha/2}}.$$

[ ]

Proposition B10D: Let $f(\lambda)$ have a Hölder-continuous third derivative on the closed unit interval, with Hölder exponent α ($0 < \alpha \leq 1$) and Hölder constant H_3 or less. If $n \geq 6$ is an integer, the error in approximating f with $U_{n,2}(f)$ is as follows:

$$\text{abs}(f - U_{n,2}(f)) \leq \frac{M_2 + M_3}{8n^2} + 9H_3/(64n^{(3+\alpha)/2})$$

$$\leq \frac{9H_3 + 8M_2 + 8M_3}{64n^{(3+\alpha)/2}}.$$

---

[75]G.G. Lorentz, "Inequalities and saturation classes for Bernstein polynomials", 1963.

[76]Draganov, Borislav R. "On simultaneous approximation by iterated Boolean sums of Bernstein operators." Results in Mathematics 66, no. 1 (2014): 21-41.

[77]Knoop, H-B., Pottinger, P., "Ein Satz vom Korovkin-Typ für C^k-Räume", Math. Zeitschrift 148 (1976).

Proof: Again, the goal is to estimate the right-hand side of (B10C-1). But this time, a different simultaneous approximation bound is employed, namely a result from Kacsó (2002)[78], which in this case works if $n \geq \max(r+2, (r+1)r) = 6$, where $r = 2$. By that result:

$$\mathrm{abs}((B_n(f))^{(2)} - f^{(2)}) \leq \frac{r(r-1)}{2n}M_2 + \frac{rM_3}{2n} + \frac{9}{8}\omega_2(f^{(2)}, 1/n^{1/2})$$

$$\leq \frac{1}{n}M_2 + M_3/n + \frac{9}{8}H_3/n^{(1+\alpha)/2},$$

where $r = 2$, using properties of ω_2, the smallest second-order modulus of continuity of $f^{(2)}$, given, for example, in Stancu et al. (2001)[79] and Gonska (1985)[^60]: if $f^{(3)}$ is Hölder continuous and $h > 0$, then $\omega_2(f^{(2)}, h) \leq h \cdot \omega_1(f^{(3)}, h) \leq h \cdot H_3 h^\alpha = H_3 h^{1+\alpha}$. Therefore—

$$\mathrm{abs}((B_n(f) - f)(B_n(f) - f)) \leq \frac{1}{8n}\left(\frac{1}{n}M_2 + M_3/n + \frac{9}{8}H_3/n^{(1+\alpha)/2}\right) \qquad (**)$$

$$\leq \frac{M_2 + M_3}{8n^2} + \frac{9H_3}{64n^{(3+\alpha)/2}} \leq \frac{9H_3 + 8M_2 + 8M_3}{64n^{(3+\alpha)/2}}.$$

[ ]

The following error bounds follow from results of Sevy (1991)[80], especially theorems 3.1 and 3.7 there:

If $f(\lambda)$ on the closed unit interval:	Then the following polynomial:	Is close to f with the following error bound:
Has continuous second derivative.	$U_{n,2}$.	$\frac{25M_2}{16n}$.
Has continuous third derivative.	$U_{n,3}$.	$\frac{25M_2}{16 \cdot n^2} + \frac{125M_3}{64n^{3/2}}$.
Has continuous fourth derivative.	$U_{n,2}$.	$\frac{195M_2}{32n^2} + \frac{45M_2}{8n^{\frac{5}{2}}} + \frac{377M_4}{128n^2} + \frac{9M_4}{8n^3} + \frac{63M_4}{16n^{\frac{5}{2}}}$.
Has continuous sixth derivative.	$U_{n,3}$.	$\frac{8775M_2}{32n^3} + \frac{2025M_2}{8n^{\frac{7}{2}}} + \frac{21489M_4}{128n^3} + \frac{513M_4}{8n^4} +$
		$\frac{3591M_4}{16n^{\frac{7}{2}}} + \frac{16965M_6}{512n^3} + \frac{783M_6}{32n^4} + \frac{6801M_6}{128n^{\frac{7}{2}}} +$
		$\frac{27M_6}{8n^{\frac{9}{2}}}$.

Lemma B11: Let $f(\lambda)$ have a continuous k-th derivative on the closed unit interval, where k is a positive whole number or zero. Let $W_n(f) = 2f - B_n(f)$. Then, for every integer $n \geq 1$:

$$\mathrm{abs}(W_n^{(k)}(f)(\lambda)) \leq 3\max_\lambda(\mathrm{abs}(f^{(k)}(\lambda))).$$

Proof: W_n can be rewritten as $f + (f - B_n(f))$, given that Bernstein polynomials are linear operators. Then the left-hand side is no more than $\max_\lambda(\mathrm{abs}(f^{(k)}(\lambda))) + \max_\lambda(\mathrm{abs}(f - B_n^{(k)}(f))(\lambda))$. By Lemma 3.6 of Sevy (1991)[81], the second term is no more than $2\max_\lambda(\mathrm{abs}(f^{(k)}(\lambda)))$, which gives the desired result. (See also a similar result of estimating the second derivative of $B_n(f)$ in Lemma 4 of Berens and Lorentz (1972)[82].) [ ]

[78]Kacsó, D.P., "Simultaneous approximation by almost convex operators", 2002.

[79]Stancu, D.D., Agratini, O., et al. Analiză Numerică şi Teoria Aproximării, 2001.

[80]Sevy, J., "Acceleration of convergence of sequences of simultaneous approximants", dissertation, Drexel University, 1991.

[81]Sevy, J., "Acceleration of convergence of sequences of simultaneous approximants", dissertation, Drexel University, 1991.

[82]Berens, H., Lorentz, G.G., "Inverse theorems for Bernstein polynomials", Indiana University Mathematics Journal 21 (1972).

Han's (2000)[83] so-called "multi-node expansions" transform certain operations that preserve all polynomials of degree up to m to those that do so up to $m + r$. For the Bernstein polynomials, this expansion is $H_{n,r}(\lambda) = B_n(f)(\lambda) + \sum_{j=1}^{r} \frac{r+1-j}{(r+1)(j!)} B_n((\lambda - t)^j f^{(j)}(t))(\lambda)$, where t is a variable that is sampled to form the Bernstein polynomial within the sum. In this case, the expansion is a polynomial of degree $n + r$ and preserves all polynomials of degree up to $r + 1$. Using Theorem 3 of Han's paper for the Bernstein case of $H_{n,r}$ gives…

Proposition B12: Let $f(\lambda)$ have a continuous r-th derivative on the closed unit interval, where $r \geq 3$ is an integer. Then, for every $n \geq 1$, $H_{n,r-2}$ is within $\frac{(r-2)!}{((r-1)!)\cdot(r!)} M_r \mu_{n,r}/n^{r/2}$ of f, where $\mu_{n,r}$ is as in Lemma B7.

> **Note:** $\mu_{n,r}/n^{r/2}$ is an upper bound on the r-th central absolute moment of X/n, discussed earlier in this section. In the special case $n = 1$ and $r = 4$, Han proved the error bound $M_4/864$ relying on a tighter bound on this moment.

10.2   Chebyshev Interpolants

The following is a method that employs *Chebyshev interpolants* to compute the Bernstein coefficients of a polynomial that comes within ϵ of $f(\lambda)$, as long as f meets certain conditions. Because the method introduces a trigonometric function (the cosine function), it appears here in the appendix and it runs too slowly for real-time or "online" use; rather, this method is more suitable for pregenerating ("offline") the approximate version of a function known in advance.

- f must be continuous on the interval $[a, b]$ and must have an r-th derivative of *bounded variation*, as described later.

- Suppose f's domain is the interval $[a, b]$. Then the *Chebyshev interpolant* of degree n of f (Wang 2023)[84], (Trefethen 2013)[85] is—

$$p(\lambda) = \sum_{k=0}^{n} c_k T_k(2\frac{\lambda - a}{b - a} - 1),$$

  where—

  - $c_k = \sigma(k,n)\frac{2}{n}\sum_{j=0}^{n}\sigma(j,n)f(\gamma(j,n))T_k(\cos(j\pi/n))$,
  - $\gamma(j,n) = a + (b - a)(\cos(j\pi/n) + 1)/2$,
  - $\sigma(k,n)$ is $1/2$ if k is 0 or n, and 1 otherwise, and
  - $T_k(x)$ is the k-th **Chebyshev polynomial of the first kind**[86] (`chebyshevt(k,x)` in the SymPy computer algebra library).

- Let $r \geq 1$ and $n > r$ be integers. If f is defined on the interval $[a, b]$, has a Lipschitz-continuous $(r-1)$-th derivative, and has an r-th derivative of *bounded variation*, then the degree-n Chebyshev interpolant of f is within $\left(\frac{(b-a)}{2}\right)^r \frac{4V}{\pi r(n-r)^r}$ of f, where V is the r-th derivative's *total variation* or greater. This relies on a theorem in chapter 7 of Trefethen (2013)[87] as well as a statement in note 1 at the end of this section.

  - If the r-th derivative is nowhere decreasing or nowhere increasing on the interval $[a, b]$, then V can equal abs($f(b) - f(a)$).

[83]Han, Xuli. "Multi-node higher order expansions of a function." Journal of Approximation Theory 124.2 (2003): 242-253. **https://doi.org/10.1016/j.jat.2003.08.001**

[84]H. Wang, "**Analysis of error localization of Chebyshev spectral approximations**", arXiv:2106.03456v3 [math.NA], 2023. https://arxiv.org/abs/2106.03456v3

[85]Trefethen, L.N., *Approximation Theory and Approximation Practice*, 2013. https://www.chebfun.org/ATAP/

[86]https://mathworld.wolfram.com/ChebyshevPolynomialoftheFirstKind.html

[87]Trefethen, L.N., *Approximation Theory and Approximation Practice*, 2013. https://www.chebfun.org/ATAP/

– If the r-th derivative is Lipschitz continuous with Lipschitz constant M or less, then V can equal $M \cdot (b - a)$ (Kannan and Kreuger 1996)[88].

– The required degree is thus $n = \text{ceil}(r + \frac{(b-a)}{2}(4V/(\pi r \epsilon))^{1/r}) \leq \text{ceil}(r + \frac{(b-a)}{2}(1.2733V/(r\epsilon))^{1/r})$, where $\epsilon > 0$ is the desired error tolerance.

- If f is so "smooth" to be *analytic* (see note 4 below) at every point in the interval $[a, b]$, a better error bound is possible, but describing it requires ideas from complex analysis that are too advanced for this article. See chapter 8 of Trefethen (2013)[89].

---

1. Compute the required degree n as given earlier, with error tolerance $\epsilon/2$.
2. Compute the values c_k as given earlier, which relate to f's Chebyshev interpolant of degree n. There will be n plus one of these values, labeled $c_0, ..., c_n$.
3. Compute the $(n+1) \times (n+1)$ matrix M described in Theorem 1 of Rababah (2003)[90].
4. Multiply the matrix by the transposed vector of values $(c_0, ..., c_n)$ to get the polynomial's Bernstein coefficients $b_0, ..., b_n$. (Transposing means turning columns to rows and vice versa.)
5. (Rounding.) For each i, replace the Bernstein coefficient b_i with $\text{floor}(b_i/(\epsilon/2) + 1/2) \cdot (\epsilon/2)$.
6. Return the Bernstein coefficients $b_0, ..., b_n$.

   Notes:

   1. The following statement can be shown. Let $f(x)$ have a Lipschitz-continuous $(r-1)$-th derivative on the interval $[a, b]$, where $r \geq 1$. If the r-th derivative of f has total variation V, then the r-th derivative of $g(x)$, where $g(x) = f(a + (b-a)(x+1)/2)$, has total variation $V \left(\frac{b-a}{2} \right)^r$ on the interval $[-1, 1]$.

   2. The method in this section doesn't require $f(\lambda)$ to have a particular minimum or maximum. If f must map the closed unit interval to itself and the Bernstein coefficients must lie on that interval, the following changes to the method are needed:

      - $f(\lambda)$ must be continuous on the closed unit interval ($a = 0$, $b = 1$) and take on only values in that interval.
      - If any Bernstein coefficient returned by the method is less than 0 or greater than 1, double the value of n and repeat the method starting at step 2 until that condition is no longer true.

   3. It would be of interest to build Chebyshev-like interpolants that sample $f(\lambda)$ at *rational* values of λ that get closer to the Chebyshev points (for example, $\cos(j\pi/n)$) with increasing n, and to find results that provide explicit bounds (with no hidden constants) on the approximation error that are close to those for Chebyshev interpolants.

   4. A function $f(x)$ is *analytic* at a point z if there is a positive number r such that f is writable as—

   $$f(x) = f(z) + f^{(1)}(z)(\lambda - z)^1/1! + f^{(2)}(z)(\lambda - z)^2/2! + ...,$$

   for every point λ satisfying $\text{abs}(\lambda - z) < r$, where $f^{(i)}$ is the i-th derivative of f. The largest value of r that makes f analytic at z is the *radius of convergence* of f at z.

---

[88]R. Kannan and C.K. Kreuger, *Advanced Analysis on the Real Line*, 1996.

[89]Trefethen, L.N., ***Approximation Theory and Approximation Practice***, 2013. https://www.chebfun.org/ATAP/

[90]Rababah, Abedallah. "**Transformation of Chebyshev–Bernstein polynomial basis**." Computational Methods in Applied Mathematics 3.4 (2003): 608-622. https://www.degruyter.com/document/doi/10.2478/cmam-2003-0038/html

10.3 Results on Derivative Bounds

These results relate to bounds on a function's derivatives. Though not yet used in this article, the results may be of interest to readers.

Proposition X1: Let $f(\lambda)$ map the closed unit interval to itself, have a maximum of m, and have a Lipschitz-continuous derivative with Lipschitz constant M. Then the derivative's absolute value is no more than—

$$\max(4m + M/4, 2\sqrt{mM}) \leq \max(4 + M, 2\sqrt{M}).$$

Proof: This is a corollary to Theorem 3.1 found in Niculescu and Buşe (2003)[91].

Proposition X2: Let $f(\lambda)$ map the closed unit interval to itself, have a maximum of m, and have a Lipschitz-continuous derivative with Lipschitz constant M (see "Definitions"). Then the derivative's absolute value is no more than $4m + M/4$ if $m/M \geq 1/16$, or $2\sqrt{mM}$ otherwise.

Proof: This is a corollary to Theorem 3.1 found in Niculescu and Buşe (2003)[92].

In the following results, denote the maximum absolute value of f's r-th derivative as $MX(f, r)$.

Proposition X4: Let $f(\lambda)$ map the closed unit interval to itself and have a Lipschitz-continuous r-th derivative for some $r \geq 4$. Then $MX(f, r-1) \leq 4^{r-1}(r!)MX(f, 0) + MX(f, r+1)/2$.

Proof: See Babenko et al. (1995)[93].

Corollary X4: Let $f(\lambda)$ map the closed unit interval to itself and have a Lipschitz-continuous fourth derivative. Then $MX(f, 3) \leq 1536 MX(f, 0) + MX(f, 5)/2$.

Corollary X5: Let $f(\lambda)$ map the closed unit interval to itself and have a Lipschitz-continuous fifth derivative. Then $MX(f, 4) \leq 30720 MX(f, 0) + MX(f, 6)/2$.

11 License

Any copyright to this page is released to the Public Domain. In case this is not possible, this page is also licensed under **Creative Commons Zero**[94].

---

[91] Niculescu, Constantin P., and Constantin Buşe. "The Hardy-Landau-Littlewood inequalities with less smoothness." J. Inequal. in Pure and Appl. Math 4 (2003).

[92] Niculescu, Constantin P., and Constantin Buşe. "The Hardy-Landau-Littlewood inequalities with less smoothness." J. Inequal. in Pure and Appl. Math 4 (2003).

[93] Babenko, V. F., V. A. Kofanov, and S. A. Pichugov. "On inequalities for norms of intermediate derivatives on a finite interval", Ukrainian Mathematical Journal 47, no. 1 (1995): 121-124.

[94] https://creativecommons.org/publicdomain/zero/1.0/