Approximations in Bernstein Form

Peter Occil

This version of the document is dated 2025-02-17.

Peter Occil

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 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 Contents

2 About This Document

This is an open-source document; for an updated version, see the source code[1] or its rendering on GitHub[2]. You can send comments on this document on the GitHub issues page[3], 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**.

3 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**[4]".

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

$$p(\lambda) = {n \choose 0}\lambda^0(1-\lambda)^{n-0}a[0] + {n \choose 1}\lambda^1(1-\lambda)^{n-1}a[1] + ... + {n \choose n}\lambda^n(1-\lambda)^{n-n}a[n],$$

where n is the polynomial's *degree* and $a[0]$, $a[1]$, ..., $a[n]$ are its n plus one *Bernstein coefficients* (which this document may simply call *coefficients* if the meaning is obvious from the context).[5]

4 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.

4.1 Approximations on the Closed Unit Interval

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

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

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

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

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

[5] choose(n, k) = (1*2*3*...*n)/((1*...*k)*(1*...*($n-k$))) = $n!/(k! * (n-k)!)$ = ${n \choose k}$ 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, $(0!) = 1$.

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**[6]. 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)[7], Guan (2009)[8], Güntürk and Li (2021, sec. 3.3)[9]. 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.
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)[10], Butzer (1955)[11]. $n \ge 6$ must be even. Evaluates f at $n/2+1$ evenly spaced points.

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

[7] 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

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

[9] 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

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

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

Name	Polynomial	Its Bernstein coefficients are found as follows:	Notes
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)[12]. $n \ge 4$ must be divisible by 4. Evaluates f at $n/2+1$ evenly spaced points.
Lorentz operator (order 2).	$Q_{n-2,2} = B_{n-2}(f) - x(1-x) \cdot B_{n-2}(f'')/(2(n-2))$.	**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 $z = (((f''((j-1)/(n-2)))\ /\ (4(n-2))) \cdot 2j(n-j)/((n-1) \cdot (n))$. The final Bernstein coefficients are now $a[0]$, ..., $a[n]$.	Holtz et al. (2011)[13]; Bernstein (1932)[14]; Lorentz (1966)[15]. $n \ge 4$; f'' is the second derivative of f. Evaluates f and f'' 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.

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

[13] 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

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

[15] 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 Hölder continuous second derivative (see "Definitions").	$U_{n,2}(f)$.	$\varepsilon = (5H_2 + 4M_2)$ / $(32n^{1+\alpha/2})$.	n=max(3, ceil((($5H_2 + 4M_2$) / $(32\epsilon))^{2/(2+\alpha)}$)).	$n \geq 3$. $0 < \alpha \leq 1$ 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, ceil((($5L_2 + 4M_2$) / $(32\epsilon))^{2/3}$)).	$n \geq 3$. Special case of previous entry.
Has Lipschitz continuous second derivative.	$Q_{n-2,2}(f)$.	ε = 0.098585 $L_2/((n-2)^{3/2})$.	n=max(4, ceil(((0.098585L_2) / $(\epsilon))^{2/3} + 2$)).	$n \geq 4$. See Proposition B10A in appendix.
Has continuous third derivative.	$L_{2,n/2}(f)$.	ε = (3*sqrt(3 − 4/n)/4)*M_3/n^2 < (3*sqrt(3)/4)*M_3/n^2 < 1.29904*M_3/n^2 ≤ 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(((1.29904M_3) / $\epsilon)^{2/3}$)). (If n is now odd, add 1.)	Tachev (2022)[16]. $n \geq 6$ must be even.
Has Hölder continuous third derivative.	$U_{n,2}(f)$.	ε = $(9H_3 + 8M_2 + 8M_3)$ / $(64n^{(3+\alpha)/2})$.	n=max(6, ceil((($9H_3 + 8M_2 + 8M_3$) / $(64\epsilon))^{2/(3+\alpha)}$)).	$n \geq 6$. $0 < \alpha \leq 1$ is third derivative's Hölder exponent. See Proposition B10D in appendix.
Has Lipschitz continuous third derivative.	$U_{n,2}(f)$.	ε = $(9H_3 + 8M_2 + 8M_3)$ / $(64n^2)$.	n=max(6, ceil((($9H_3 + 8M_2 + 8M_3$) / $(64\epsilon))^{1/2}$)).	$n \geq 6$. 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(2)/4) * sqrt(L_3/ ε))) ≤ max(4,ceil((35356/100000) * sqrt(L_3/ ε))). (Round n up to nearest multiple of 4.)	$n \geq 4$ must be divisible by 4. See Proposition B10 in appendix.
Has Lipschitz continuous derivative.	$B_n(f)$.	$\varepsilon = L_1/(8{*}n)$.	n = ceil(L_1/(8* ε)).	Lorentz (1963)[17].[18]

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

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

[18] 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).

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 derivative.	$B_n(f)$.	$\varepsilon = H_1/(4*n^{(1+\alpha)/2})$.	n = ceil$((H_1/(4*\varepsilon))^{2/(1+\alpha)})$.	Schurer and Steutel (1975)[19]. $0 < \alpha \leq 1$ is derivative's Hölder exponent.
Is Hölder continuous.	$B_n(f)$.	$\varepsilon = H_0*(1/(4*n))^{\alpha/2}$.	n = ceil$((H_0/\varepsilon)^{2/\alpha}/4)$.	Kac (1938)[20]. $0 < \alpha \leq 1$ is f's Hölder exponent.
Is Lipschitz continuous.	$B_n(f)$.	$\varepsilon = L_0$*sqrt$(1/(4*n))$.	n = 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$((L_0*1.08989/\varepsilon)^2)$.	Sikkema (1961)[21].

Note: In addition, by analyzing the proof of Theorem 2.4 of Güntürk and Li (2021, sec. 3.3)[22], the following error bounds for $U_{n,3}$ *appear* to be true:

- If $f(\lambda)$ has continuous fifth derivative: $\varepsilon = 4.0421*\max(M_0,...,M_5)/n^{5/2}$.
- If $f(\lambda)$ has continuous sixth derivative: $\varepsilon = 4.8457*\max(M_0,...,M_6)/n^3$.

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—

[19] Schurer and Steutel, "On an inequality of Lorentz in the theory of Bernstein polynomials", 1975.
[20] Kac, M., "Une remarque sur les polynômes de M. S. Bernstein", Studia Math. 7, 1938.
[21] Sikkema, P.C., "Der Wert einiger Konstanten in der Theorie der Approximation mit Bernstein-Polynomen", 1961.
[22] 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
[23] E. Voronovskaya, "Détermination de la forme asymptotique d'approximation des fonctions par les polynômes de M. Bernstein", 1932.

- 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 $B \ge 1$, $l \ge 1$, $\gamma \ge 1$ such that f's n-th derivative's absolute value is not greater than $Bl^n n^{\gamma n}$ for every $n \ge 1$ (Kawamura et al. 2015)[24]; see also (Gevrey 1918)[25]. In this case, for each $n \ge 1$—

   - 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).

4.2 Taylor Polynomials for "Smooth" Functions

If $f(\lambda)$ is "smooth" enough on the closed unit interval and if ϵ is big enough, then Taylor's theorem shows how to build a polynomial 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 $n \ge 0$ 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 $\epsilon \le f(0) \le 1-\epsilon$ and $\epsilon \le f(1) \le 1-\epsilon$, 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 $\epsilon \ge M_{n+1}/((n+1)!)$, and
5. $f(0)$ is known as well as $f^{(1)}(0), ..., f^{(n)}(0)$.

Then the n-th *Taylor polynomial* centered at 0, given later, is within ϵ of f:

$$P(\lambda) = a_0\lambda^0 + a_1\lambda^1 + ... + a_n\lambda^n,$$

where $a_0 = f(0)$ and $a_i = f^{(i)}(0)/(i!)$ for $i \ge 1$.

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 $f(\lambda) = c_0\lambda^0 + c_1\lambda^1 + ... + c_i\lambda^i + ...$ whenever $0 \le \lambda \le 1$ (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, ...)), and
> - meet the so-called "ratio test",

[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.

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 P, when written in its "power" form or "monomial" form, has "power" coefficients $a_0, ..., a_n$.

Now, rewrite $P(\lambda)$ as a polynomial in Bernstein form. (See "**Computational Issues**" for details.) Let $b_0, ..., b_n$ 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.

4.3 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—

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

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 $[0, 1]$.

[26] Carvalho, Luiz Max, and Guido A. Moreira. "**Adaptive truncation of infinite sums: applications to Statistics**", arXiv:2202.06121 (2022). https://arxiv.org/abs/2202.06121

[27] Bărbosu, D., "The Bernstein operators on any finite interval revisited", Creat. Math. Inform. 20 (2020).

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 $[1, 3]$.

4.4 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.

4.5 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.

---

[28] Tsai, Y., Farouki, R.T., "Algorithm 812: BPOLY: An Object- Oriented Library of Numerical Algorithms for Polynomials in Bernstein Form", ACM Transactions on Mathematical Software, June 2001.

[29] As an example, Mastroianni and Occorsio (1977) approximate an integral this way using iterated Boolean sums of Bernstein polynomials (which include $U_{n,2}$). G. Mastroianni, M.R. Occorsio, "Una generalizzazione dell'operatore di Bernstein", 1977.

[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.

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 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.	$B_n(f)$.	$\epsilon = rM_r(r-1)/(2n) + 5H_r/(4n^{\alpha/2}) \le (rM_r(r-1)/2 + 5H_r/4)/n^{\alpha/2}$.	$n=\text{ceil}(\max(r+1, \left(\frac{(5H_r+2M_r(r^2-r))^2}{16\epsilon^2}\right)^{1/\alpha}))$.	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]).

4.6 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].

[33]Tsai, Y., Farouki, R.T., "Algorithm 812: BPOLY: An Object- Oriented Library of Numerical Algorithms for Polynomials in Bernstein Form", ACM Transactions on Mathematical Software, June 2001.

[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, Y., Farouki, R.T., "Algorithm 812: BPOLY: An Object- Oriented Library of Numerical Algorithms for Polynomials in Bernstein Form", ACM Transactions on Mathematical Software, June 2001.

- 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).

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.

5 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 has no root that is a complex number whose real part is between 0 and 1 (Borwein 1979, theorem 29)[43].

6 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.

[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

[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).

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).

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.
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**[44].

7 Notes

8 Appendix

8.1 Results Used in Approximations by Polynomials

Lemma A1: Let—

$$f(x) = a_0x^0 + a_1x^1 + ...,$$

where the a_i are constants each 0 or greater and sum to a finite value and where $0 \le x \le 1$ (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 $x \ge 0$ (and thus wherever $0 \le x \le 1$). Each of its terms has a maximum at 1 since—

- for $n = 0$, $a_0x^0 = a_0$ is a nonnegative constant (which trivially reaches its maximum at 1), and
- for each n where $a_0 = 0$, a_0x^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—

$$f'(x) = 1 \cdot a_1x^0 + ... + i \cdot a_ix^{i-1} + ...,$$

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 $x \ge 0$, and thus on its domain, f is convex, which completes the proof. [ ]

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

$$g_n(x) = a_0x^0 + ... + a_nx^n,$$

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

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

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 1. The same is true for $f - g_n$, consisting of the remaining terms of f. Since the latter has a maximum at 1, the maximum error is $\epsilon = f(1) - g_n(1)$. [ ]

For a function f 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 the point 1, so the algorithms given in Carvalho and Moreira (2022)[45] — 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 $B_n(f) \le f$ for every $n \ge 1$. 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)[46]. 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.

Proof: The assumptions on f imply that $B_n(f) \le 2f$ (Li 2000)[47], 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)[48] 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").

[45] Carvalho, Luiz Max, and Guido A. Moreira. "**Adaptive truncation of infinite sums: applications to Statistics**", arXiv:2202.06121 (2022). https://arxiv.org/abs/2202.06121

[46] 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.

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

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

- 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 *absolute moment* of X/n or the ratio of heads to tosses is denoted $M_{n,r}(p) = \mathbb{E}[\text{abs}(X/n - \mathbb{E}[X/n])^r] = B_n(\text{abs}(\lambda - p)^r)(p)$.

The following results bound the absolute value of $T_{n,r}$, $S_{n,r}$, and $M_{n,r}$.[49]

Result B4 (Molteni (2022)[50]): If r is an even integer such that $0 \le r \le 44$, then for every integer $n \ge 1$, $\text{abs}(T_{n,r}(p)) \le ((r!)/(((r/2)!)8^{r/2})) \cdot n^{r/2}$ and $\text{abs}(S_{n,r}(p)) \le ((r!)/(((r/2)!)8^{r/2})) \cdot (1/n^{r/2})$.

Result B4A (Adell et al. (2015)[51]): For every odd integer $r \ge 1$, $T_{n,r}(p)$ is positive whenever $0 \le p < 1/2$, and negative whenever $1/2 < p \le 1$.

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

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

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)[52] 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 $r = 4$.

Result B6 (Adell and Cárdenas-Morales (2018)[53]): Let $\sigma(r,t) = (r!)/(((r/2)!)t^{r/2})$. If $r \ge 0$ is an even integer, then—

- for every integer $n \ge 1$, $\text{abs}(T_{n,r}(p)) \le \sigma(r,6)n^{r/2}$ and $\text{abs}(S_{n,r}(p)) \le \sigma(r,6)/n^{r/2}$, and
- for every integer $n \ge 1$, $\text{abs}(T_{n,r}(1/2)) \le \sigma(r,8)n^{r/2}$ and $\text{abs}(S_{n,r}(1/2)) \le \sigma(r,8)/n^{r/2}$.

Lemma B9: Let $f(\lambda)$ have a Lipschitz continuous r-th derivative on the closed unit interval (see "**Definitions**"), where $r \ge 0$ is an integer, and let M be equal to or greater than the r-th derivative's Lipschitz constant. Denote $B_n(f)$ as the Bernstein polynomial of f of degree n. Then, for every $0 \le x_0 \le 1$:

1. f can be written as $f(\lambda) = R_{f,r}(\lambda, x_0) + f(x_0) + \sum_{i=1}^{r}(\lambda - x_0)^i f^{(i)}(x_0)/(i!)$ where $f^{(i)}$ is the i-th derivative of f.
2. If r is odd, $\text{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \le M/((((r+1)/2)!)(\beta n)^{(r+1)/2})$ for every integer $n \ge 1$, where β is 8 if $r \le 43$ and 6 otherwise.
3. If $r = 0$, $\text{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \le M/(2n^{1/2})$ for every integer $n \ge 1$.
4. If r is even and greater than 0, $\text{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \le \frac{M}{(r+1)!n^{(r+1)/2}}\left(\frac{2\cdot(r+1)!(r)!}{\gamma^{r+1}((r/2)!)^2}\right)^{1/2}$ for every integer $n \ge 2$, where γ is 8 if $r \le 42$ and 6 otherwise.

[49] *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 $g(0) + g(1) + \ldots + g(n)$, and $\sum_{k\ge 0} g(k)$ means $g(0) + g(1) + \ldots$.

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

[51] 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.

[52] 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.

[53] 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

Proof: The well-known result of part 1 says f equals the *Taylor polynomial* of degree r at x_0 plus the *Lagrange remainder*, $R_{f,r}(\lambda, x_0)$. A result found in Gonska et al. (2006)[54], which applies for any integer $r \geq 0$, bounds that Lagrange remainder [55]. By that result, because f's r-th derivative is Lipschitz continuous—

$$\text{abs}(R_{f,r}(\lambda, x_0)) \leq \frac{\text{abs}(\lambda - x_0)^r}{r!} M \frac{\text{abs}(\lambda - x_0)}{r+1} = M \frac{\text{abs}(\lambda - x_0)^{r+1}}{(r+1)!}.$$

The goal is now to bound the Bernstein polynomial of $\text{abs}(\lambda - x_0)^{r+1}$. This is easiest to do if r is odd.

If r is odd, then $(\lambda - x_0)^{r+1} = \text{abs}(\lambda - x_0)^{r+1}$, so by Results B4 and B6, the Bernstein polynomial of $\text{abs}(\lambda - x_0)^{r+1}$ can be bounded as follows:

$$\text{abs}(B_n((\lambda - x_0)^{r+1})(x_0)) = \text{abs}(S_{n,r+1}(x_0)) \leq \frac{(r+1)!}{(((r+1)/2)!)\beta^{(r+1)/2}} \frac{1}{n^{(r+1)/2}} = \sigma(r+1, n),$$

where β is 8 if $r \leq 43$ and 6 otherwise. Therefore—

$$\text{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \leq \frac{M}{(r+1)!} \text{abs}(B_n((\lambda - x_0)^{r+1})(x_0))$$

$$\leq \frac{M}{(r+1)!} \frac{(r+1)!}{(((r+1)/2)!)\beta^{(r+1)/2}} \frac{1}{n^{(r+1)/2}} = \frac{M}{(((r+1)/2)!)(\beta n)^{(r+1)/2}}.$$

If r is 0, then the Bernstein polynomial of $\text{abs}(\lambda - x_0)^1$ is bounded by $\sqrt{x_0(1-x_0)/n}$ for every integer $n \geq 1$ (Cheng 1983)[56], so—

$$\text{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \leq \frac{M}{(r+1)!} \sqrt{x_0(1-x_0)/n} \leq \frac{M}{(r+1)!} \frac{1}{2n^{1/2}} = \frac{M}{2n^{1/2}}.$$

If r is even and greater than 0, the Bernstein polynomial for $\text{abs}(\lambda - x_0)^{r+1}$ can be bounded as follows for every $n \geq 2$, using **Schwarz's inequality**[57] (see also Bojanic and Shisha [1975][58] for the case $r = 4$):

$$B_n(\text{abs}(\lambda - x_0)^{r+1})(x_0) = B_n((\text{abs}(\lambda - x_0)^{r/2} \text{abs}(\lambda - x_0)^{(r+2)/2})^2)(x_0)$$

$$\leq \sqrt{\text{abs}(S_{n,r}(x_0))} \sqrt{\text{abs}(S_{n,r+2}(x_0))} \leq \sqrt{\sigma(r, n)} \sqrt{\sigma(r+2, n)}$$

$$\leq \frac{1}{n^{(r+1)/2}} \left(\frac{2 \cdot (r+1)!(r)!}{\gamma^{r+1}((r/2)!)^2} \right)^{1/2},$$

where γ is 8 if $r \leq 42$ and 6 otherwise. Therefore—

[54] Gonska, H.H., Piţul, P., Raşa, I., "On Peano's form of the Taylor remainder, Voronovskaja's theorem and the commutator of positive linear operators", In Numerical Analysis and Approximation Theory, 2006.

[55] The result from Gonska et al. actually applies if the r-th derivative belongs to a broader class of continuous functions than Lipschitz continuous functions, but this feature is not used in this proof.

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

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

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

$$\text{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \le \frac{M}{(r+1)! \cdot n^{(r+1)/2}} \left(\frac{2 \cdot (r+1)!(r)!}{\gamma^{r+1}((r/2)!)^2}\right)^{1/2}.$$

[ ]

Notes:

1. If a function $f(\lambda)$ has a continuous r-th derivative on its domain (where $r \ge 0$ is an integer), then by Taylor's theorem for real variables, $R_{f,r}(\lambda, x_0)$, is writable as $f^{(r)}(c) \cdot (\lambda - x_0)^r/(r!)$, for some c between λ and x_0 (and thus on f's domain) (DLMF [59] **equation 1.4.36**[60]). Thus, by this estimate, $\text{abs}(R_{f,r}(\lambda, x_0)) \le \frac{M}{r!}(\lambda - x_0)^r$.
2. It would be interesting to strengthen this lemma, at least for $r \le 10$, 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 $\text{abs}(\lambda - x_0)^{r+1}$ can be bounded in this way (Lorentz 1966)[61].

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 $0 \le x_0 \le 1$:

If r is:	Then $\text{abs}(B_n(R_{f,r}(\lambda, x_0))(x_0)) \le$ ...
0.	$M/(2n^{1/2})$ for every integer $n \ge 1$.
0.	$M \cdot \sqrt{x_0(1-x_0)/n}$ for every integer $n \ge 1$.
1.	$M/(8n)$ for every integer $n \ge 1$.
2.	$\sqrt{3}M/(48n^{3/2}) < 0.03609M/n^{3/2}$ for every integer $n \ge 2$.
3.	$M/(128n^2)$ for every integer $n \ge 1$.
4.	$\sqrt{5}M/(1280n^{5/2}) < 0.001747M/n^{5/2}$ for every integer $n \ge 2$.
5.	$M/(3072n^3)$ for every integer $n \ge 1$.

Proposition B10: Let $f(\lambda)$ have a Lipschitz continuous third derivative on the closed unit interval. For each $n \ge 4$ 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 $\Lambda_3/(8n^2)$ 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)[62].

Because f has a Lipschitz continuous third derivative, f has the Lagrange 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, linear, and constant functions), so that $L_{3,n/4}(f) = f$ whenever f is a polynomial of degree 3 or less (Ditzian and Totik 1987)[63], Butzer (1955)[64], May (1976)[65]. Because of this, it can be assumed without loss of generality that $f(x_0) = 0$.

[59] *NIST Digital Library of Mathematical Functions*, **https://dlmf.nist.gov/** , Release 1.1.9 of 2023-03-15.
[60] https://dlmf.nist.gov/1.4.E36
[61] G.G. Lorentz, "The degree of approximation by polynomials with positive coefficients", 1966.
[62] Tachev, Gancho. "**Linear combinations of two Bernstein polynomials**", *Mathematical Foundations of Computing*, 2022. https://doi.org/10.3934/mfc.2022061
[63] Ditzian, Z., Totik, V., *Moduli of Smoothness*, 1987.
[64] Butzer, P.L., "Linear combinations of Bernstein polynomials", Canadian Journal of Mathematics 15 (1953).
[65] May, C.P., "Saturation and inverse theorems for a class of exponential-type operators", Canadian Journal of Mathematics 28 (1976).

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 $0 \le x_0 \le 1$. In turn (using Corollary B9A)—

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

$$\le (1/3)\frac{\Lambda_3}{128(n/4)^2} + 2\frac{\Lambda_3}{128(n/2)^2} + (8/3)\frac{\Lambda_3}{128n^2} = \Lambda_3/(8n^2).$$

[ ]

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

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

(where α_i are real numbers and $n(i) \ge 1$ 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)[66].

Proposition B10A: Let $f(\lambda)$ have a Lipschitz continuous second derivative on the closed unit interval. Let $Q_{n,2}(f) = B_n(f)(x) - \frac{x(1-x)}{2n} B_n(f'')(x)$ be the *Lorentz operator* of order 2 (Holtz et al. 2011)[67], (Lorentz 1966)[68], which is a polynomial in Bernstein form of degree $n+2$. Then if $n \ge 2$ is an integer, $Q_{n,2}(f)$ is within $\frac{L_2(\sqrt{3}+3)}{48n^{3/2}} < 0.098585 L_2/(n^{3/2})$ 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)[69] 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'', the second derivative of f, has the Lagrange remainder $R_{f'',0}(\lambda, x_0)$. Thus, using Corollary B9A, the error bound can be written as—

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

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

[ ]

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.

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

[67] 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

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

[69] 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)[70] 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 and max(abs(f)) means the maximum of the absolute value of the function f.

Proposition B10C: Let $f(\lambda)$ have a Hölder continuous second derivative on the closed unit interval, with Hölder exponent α ($0 < \alpha \leq 1$) and Hölder constant H_2 or less. 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:

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

where M_2 is the maximum of that second derivative's absolute value or greater.

Proof: This proof is inspired by a result in Draganov (2014, Theorem 4.1)[71].

The error to be bounded can be expressed as 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)[72], 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—

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

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

[ ]

> **Note**: The error bound $0.75M_2/n^2$ for $U_{n,2}$ is false in general if $f(\lambda)$ is assumed only to be nonnegative, concave, and have a continuous second derivative on the closed unit interval. A counterexample is $f(\lambda) = (1 - (1 - 2\lambda)^{2.5})/2$ if $\lambda < 1/2$ and $(1 - (2\lambda - 1)^{2.5})/2$ otherwise.

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:

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

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

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

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

$$\leq \frac{9H_3 + 8\max(\text{abs}(f^{(2)})) + 8\max(\text{abs}(f^{(3)}))}{64n^{(3+\alpha)/2}}.$$

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)[73], which in this case works if $n \geq \max(r+2, (r+1)r) = 6$, where $r = 2$. By that result:

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

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

where $r = 2$, $M_2 = \max(\text{abs}(f^{(2)}))$, and $M_3 = \max(\text{abs}(f^{(3)}))$, using properties of ω_2, the second-order modulus of continuity of $f^{(2)}$, given in Stancu et al. (2001)[74]. Therefore—

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

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

[ ]

In a similar way, it's possible to prove an error bound for $U_{n,3}$ that applies to functions with a Hölder continuous fourth or fifth derivative, by expressing the error bound as $\text{abs}((B_n(f)-f)((B_n(f)-f)(B_n(f)-f)))$ and replacing the values for M_2, M_3, and H_3 in the bound proved at the end of Proposition B10D with upper bounds for $\text{abs}((B_n(f))^{(2)} - f^{(2)})$, $\text{abs}((B_n(f))^{(3)} - f^{(3)})$, and $\text{abs}((B_n(f))^{(4)} - f^{(4)})$, respectively.

8.2 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)[75], (Trefethen 2013)[76] is—

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

where—

[73]Kacsó, D.P., "Simultaneous approximation by almost convex operators", 2002.
[74]Stancu, D.D., Agratini, O., et al. *Analiză Numerică şi Teoria Aproximării*, 2001.
[75]H. Wang, "**Analysis of error localization of Chebyshev spectral approximations**", arXiv:2106.03456v3 [math.NA], 2023. https://arxiv.org/abs/2106.03456v3
[76]Trefethen, L.N., ***Approximation Theory and Approximation Practice***, 2013. https://www.chebfun.org/ATAP/

- $c_k = \sigma(k,n)\frac{2}{n}\sum_{j=0}^{n}\sigma(j,n)f(\gamma(j,n))T_k(\cos(j\pi/n))$,
- $\gamma(j,n) = a + (b-a)(\cos(j\pi/n)+1)/2$,
- $\sigma(k,n)$ 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**[77] (`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)[78] 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 $\text{abs}(f(b)-f(a))$.
    - 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)[79].
    - 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)[80].

---

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)[81].
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.

[77] https://mathworld.wolfram.com/ChebyshevPolynomialoftheFirstKind.html
[78] Trefethen, L.N., ***Approximation Theory and Approximation Practice***, 2013. https://www.chebfun.org/ATAP/
[79] R. Kannan and C.K. Kreuger, *Advanced Analysis on the Real Line*, 1996.
[80] Trefethen, L.N., ***Approximation Theory and Approximation Practice***, 2013. https://www.chebfun.org/ATAP/
[81] 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

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.

9 License