Constr. Approx. (1985) 1: 333–348

CONSTRUCTIVE APPROXIMATION

On the Bernstein Conjecture in Approximation Theory

Richard S. Varga and Amos J. Carpenter

Abstract. With $E_{2n}(|x|)$ denoting the error of best uniform approximation to $|x|$ by polynomials of degree at most $2n$ on the interval $[-1, +1]$, the famous Russian mathematician S. Bernstein in 1914 established the existence of a positive constant β for which

$$\lim_{n\to\infty} 2nE_{2n}(|x|) =: \beta.$$

Moreover, by means of numerical calculations, Bernstein determined, in the same paper, the following upper and lower bounds for β:

$$0.278 < \beta < 0.286.$$

Now, the average of these bounds is 0.282, which, as Bernstein noted as a "curious coincidence," is very close to $1/(2\sqrt{\pi}) = 0.28209\,47917\ldots$. This observation has over the years become known as the Bernstein Conjecture: Is $\beta = 1/(2\sqrt{\pi})$? We show here that the Bernstein conjecture is *false*. In addition, we determine rigorous upper and lower bounds for β, and by means of the Richardson extrapolation procedure, estimate β to approximately 50 decimal places.

1. Introduction

For any real function $f(x)$ with domain $[-1, +1]$, its modulus of continuity is defined as usual by

$$(1.1)\qquad \omega(\delta; f) := \sup_{\substack{|x_1 - x_2| \le \delta \\ x_1, x_2 \in [-1, +1]}} |f(x_1) - f(x_2)|.$$

With π_n denoting the set of all real polynomials of degree at most n $(n = 0, 1, \ldots)$, the following is a well-known result of Jackson (cf. Meinardus [8, p. 56], Rivlin [10, p. 22]):

Date received: February 22, 1985. Communicated by P. Henrici.
AMS classification: 41A25.
Key words and phrases: Best uniform approximation, Polynomials, Chebyshev series, Remez algorithm.

Theorem A. *If* $f \in C[-1, +1]$, *then*

$$E_n(f) \le 6\omega\left(\frac{1}{n}; f\right) \qquad (n = 1, 2, \ldots), \tag{1.2}$$

where

$$E_n(f) := \inf\left\{\|f - g\|_{L_\infty[-1,+1]}: g \in \pi_n\right\}. \tag{1.3}$$

For the particular continuous function $|x|$ on $[-1, +1]$, it is easily seen that

$$\omega(\delta; |x|) = \delta \qquad (0 < \delta \le 1),$$

so that from (1.2) of Theorem A,

$$E_n(|x|) \le \frac{6}{n} \qquad (n = 1, 2, \ldots). \tag{1.4}$$

Since $|x|$ is an even continuous function on $[-1, +1]$, then so is its (unique) best uniform approximation from π_n on $[-1, +1]$, for any $n \ge 0$ (cf. Rivlin [10, p. 43, Exercise 1.1]). Combining this observation with the Chebyshev alternation characteristic of best uniform approximation by polynomials, it readily follows (cf. [10, p. 26]) that

$$E_{2n}(|x|) = E_{2n+1}(|x|) \qquad (n = 0, 1, \ldots). \tag{1.5}$$

Hence, it suffices for our purposes to consider only the manner in which the sequence $\{E_{2n}(|x|)\}_{n=1}^{\infty}$ decreases to zero. From (1.4), we evidently have

$$2nE_{2n}(|x|) \le 6 \qquad (n = 1, 2, \ldots). \tag{1.6}$$

Actually, on expanding $|x|$ in a Chebyshev series on $(-1, 1)$, i.e.,

$$|x| = \frac{4}{\pi}\left\{\frac{1}{2} + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}T_{2m}(x)}{(2m-1)(2m+1)}\right\}, \tag{1.7}$$

where $T_n(x)$ denotes the nth Chebyshev polynomial (of the first kind), the famous Russian mathematician S. Bernstein [2] showed in 1914 that

$$\left| |x| - \frac{4}{\pi}\left\{\frac{1}{2} + \sum_{m=1}^{n} \frac{(-1)^{m+1}T_{2m}(x)}{(2m-1)(2m+1)}\right\}\right| \le \frac{2}{\pi(2n+1)} \tag{1.8}$$

for all $x \in [-1, +1]$ and all $n \ge 1$. Since the approximation to $|x|$ in (1.8) is a particular polynomial of degree $2n$, it follows from (1.3) and (1.8) that

$$2nE_{2n}(|x|) \le \frac{4n}{\pi(2n+1)} < \frac{2}{\pi} = 0.63661\,97723\ldots, \tag{1.9}$$

which improves the upper bound of (1.6).

It turns out that (1.9) can be significantly sharpened. In his fundamental paper [2], Bernstein gave a long and difficult proof that there *exists* a positive constant, which we call β (β for "Bernstein"), such that

$$\lim_{n\to\infty} 2nE_{2n}(|x|) = \beta. \tag{1.10}$$

(We remark that a simple proof of (1.10) has yet to be found.) In addition, not being content with just the *existence* of this constant β, Bernstein, using crude calculations based on extremely ingenious methods, deduced in [2] the following rigorous upper and lower bounds for β:

$$0.278 < \beta < 0.286. \tag{1.11}$$

Moreover, Bernstein noted [2, p. 56] as a "curious coincidence" that the constant

$$\frac{1}{2\sqrt{\pi}} = 0.28209\,47917\ldots \tag{1.12}$$

also satisfies the bounds of (1.11) and, as one quickly sees, is very nearly the average of these bounds. This observation has, over the years, become known as the **Bernstein Conjecture**:

$$\beta \stackrel{?}{=} \frac{1}{2\sqrt{\pi}}. \tag{1.13}$$

In the 70 years since Bernstein's work appeared, the question of the truth of this Conjecture has remained unresolved, despite numerical attacks by several authors (cf. Bell and Shah [1], Bojanic and Elkins [3], and Salvati [11]). The reasons that this conjecture has remained open so long appear to be that (i) the determination of $E_{2n}(|x|)$ for n reasonably large is numerically *nontrivial*, and (ii) the convergence of $2nE_{2n}(|x|)$ to β [cf. (1.10)] is quite *slow*.

The main results of this paper are the following:

I. In Section 2, the numbers $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ are determined with accuracies of nearly 95 decimal digits.

II. In Section 3, upper bounds $\{2\mu_m\}_{m=0}^{100}$ for β [cf. (1.10)] are determined, where

$$\beta \le 2\mu_{100} < 2\mu_{99} < \ldots < 2\mu_0. \tag{1.14}$$

III. In Section 4, lower bounds $\{l_m\}_{m=1}^{20}$ for β are determined, where

$$l_1 < l_2 < \ldots < l_{20} \le \beta. \tag{1.15}$$

IV. In Section 5, Richardson extrapolation (with $x_n = 1/n^2$) of the numbers $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ of result I gives an approximation of β that is probably accurate to 50 decimal places.

Based on results I, II, and III, it follows that

$$l_{20} = 0.28016\,85460\ldots \le \beta \le 0.28017\,33791\ldots = 2\mu_{100}, \tag{1.16}$$

which, from (1.12), is sufficient to *disprove* the Bernstein Conjecture of (1.13). We further see from the bounds of (1.16) that β, when rounded to five decimal places, is exactly 0.28017. Similarly, the bounds of (1.16) give

$$\beta = 0.280171 + \delta, \qquad \text{where } |\delta| < 2.5 \cdot 10^{-6}. \tag{1.17}$$

Actually, the numerical results of Section 5 similarly indicate that, to 50 decimal places,

$$\beta \doteqdot 0.28016\,94990\,23869\,13303\,64364\,91230\,67200\,00424\,82139\,81236. \tag{1.18}$$

It remains an open question whether the constant β, as approximated in (1.18), can be expressed in terms of known mathematical functions and/or constants.

Finally, based on our extrapolations of Section 5, we make the following new

Conjecture. $2nE_{2n}(|x|)$ *admits an asymptotic expansion of the form*

$$2nE_{2n}(|x|) \stackrel{?}{=} \beta - \frac{K_1}{n^2} + \frac{K_2}{n^4} - \frac{K_3}{n^6} + \ldots, \qquad (n \to \infty), \tag{1.19}$$

where the constants K_j *(independent of* n*) are all positive.*

The approximate values of $\{K_j\}_{j=1}^{10}$, determined from Richardson extrapolations of $\{2nE_{2n}(|x|)\}_{n=1}^{52}$, are given in Section 5. We hope to attack the above conjecture in the future.

2. Computing the Numbers $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ with High Accuracy

Let $\hat{p}_{2n}(x)$ in π_{2n} denote the unique best uniform approximation of $|x|$ from π_{2n} on $[-1, +1]$; i.e. [cf. (1.3)],

$$\| |x| - \hat{p}_{2n}(x) \|_{L_\infty[-1,+1]} = E_{2n}(|x|) \qquad (n = 1, 2, \ldots). \tag{2.1}$$

As mentioned in Section 1, because $|x|$ is even on $[-1, +1]$, so is its best approximation $\hat{p}_{2n}(x)$, so that

$$\hat{p}_{2n}(x) = \sum_{j=0}^{n} a_j(n) x^{2j} \qquad (n = 1, 2, \ldots). \tag{2.2}$$

As was similarly done by Salvati [11] in his calculations, we make the change of variables $x^2 = t$, $t \in [0, 1]$, which changes our approximation problem to

$$E_{2n}(|x|) = \tilde{E}_n(\sqrt{t}\,; [0,1]) := \inf\left\{ \|\sqrt{t} - h_n(t)\|_{L_\infty[0,1]} : h_n \in \pi_n \right\}. \tag{2.3}$$

If we write

$$\tilde{E}_n(\sqrt{t}\,; [0,1]) = \|\sqrt{t} - \hat{h}_n(t)\|_{L_\infty[0,1]} \qquad (\text{where } \hat{h}_n \in \pi_n), \tag{2.4}$$

then clearly [cf. (2.2)]

$$\hat{p}_{2n}(x) = \hat{h}_n(x^2) \qquad (n = 1, 2, \ldots). \tag{2.5}$$

Thus, determining $E_{2n}(|x|)$ and $\hat{p}_{2n}(x)$ is equivalent to determining $\tilde{E}_n(\sqrt{t}\,; [0,1])$ and $\hat{h}_n(t)$.

The minimization problem (2.3) was solved using the following essentially standard implementation of the (second) Remez algorithm (cf. [8, p. 105]):

Step 1. Let $S := \{t_j\}_{j=0}^{n+1}$ be a set of $n + 2$ distinct (alternation) points in $[0, 1]$ satisfying

$$0 = t_0 < t_1 < \ldots < t_{n+1} = 1. \tag{2.6}$$

Step 2. Find the unique polynomial $h_n(t)$ and the constant λ, a linear problem, such that

$$h_n(t_j) + (-1)^j \lambda = \sqrt{t_j} \qquad (j = 0, 1, \ldots, n+1). \tag{2.7}$$

Table 2.1. $\{2nE_{2n}(|x|)\}_{n=1}^{52}$.

n	$2nE_{2n}(\lvert x\rvert)$	n	$2nE_{2n}(\lvert x\rvert)$
1	0.25000 00000 00000 00000	27	0.28010 92365 22206 18525
2	0.27048 35971 11137 10107	28	0.28011 34608 89950 28384
3	0.27557 43724 01175 38604	29	0.28011 72562 49499 61792
4	0.27751 78246 75052 69646	30	0.28012 06787 72662 82833
5	0.27845 11855 35508 60152	31	0.28012 37757 31660 88450
6	0.27896 79174 64958 70636	32	0.28012 65871 38731 91844
7	0.27928 29449 58518 02460	33	0.28012 91470 43904 51720
8	0.27948 88375 94507 44771	34	0.28013 14845 70012 61069
9	0.27963 06574 10128 20125	35	0.28013 36247 44030 04676
10	0.27973 24337 71973 82968	36	0.28013 55891 69271 11713
11	0.27980 79172 88743 87383	37	0.28013 73965 72336 69662
12	0.27986 54321 23793 27279	38	0.28013 90632 50782 89591
13	0.27991 02543 15557 69036	39	0.28014 06034 41582 48218
14	0.27994 58584 85782 13247	40	0.28014 20296 25997 94087
15	0.27997 46066 86407 49231	41	0.28014 33527 83104 08169
16	0.27999 81519 56316 72827	42	0.28014 45826 01611 08707
17	0.28001 76771 33297 25379	43	0.28014 57276 57645 50097
18	0.28003 40474 14993 50964	44	0.28014 67955 64600 41624
19	0.28004 79072 85905 85156	45	0.28014 77930 99959 13546
20	0.28005 97447 60423 15265	46	0.28014 87263 13048 74446
21	0.28006 99348 31809 43067	47	0.28014 96006 16931 43684
22	0.28007 87694 75287 53423	48	0.28015 04208 67046 95023
23	0.28008 64787 57075 57049	49	0.28015 11914 28744 92326
24	0.28009 32459 38808 50547	50	0.28015 19162 35465 27355
25	0.28009 92184 52382 83558	51	0.28015 25988 39017 81632
26	0.28010 45159 86556 70489	52	0.28015 32424 53163 84249

Thus, $h_n(t)$ is the best approximation from π_n to $\sqrt{t}$ on this discrete alternation set S, with (alternating) error $|\lambda|$, so that, in analogy with the notation of (2.3),

$$\|\sqrt{t} - h_n(t)\|_{L_\infty(S)} = \tilde{E}_n(\sqrt{t}\,; S) = |\lambda|. \tag{2.8}$$

Because S is a subset of $[0, 1]$, then evidently

$$\|\sqrt{t} - h_n(t)\|_{L_\infty[0,1]} - |\lambda| \geq 0. \tag{2.9}$$

Step 3. If $\|\sqrt{t} - h_n(t)\|_{L_\infty[0,1]} - |\lambda|$ is sufficiently small, the iteration is terminated. Otherwise, find a new alternation set S from the set of local extrema in $[0, 1]$ (with alternating signs) of the function $\sqrt{t} - h_n(t)$ from the previous Step 2, and repeat Steps 2 and 3, etc.

We remark that the sequence of $|\lambda|$'s generated from repeated application of this Remez algorithm is monotone increasing.

Starting with the particular alternation set $S^{(0)} := \{t_j^{(0)}\}_{j=0}^{n+1}$, where

$$t_{n+1-j}^{(0)} := \frac{1}{2}\left\{1 + \cos\left(\frac{j\pi}{n+1}\right)\right\} \qquad (j = 0, 1, \ldots, n+1) \tag{2.10}$$

are the $n + 2$ extreme points of the Chebyshev polynomial $T_{n+1}(2t - 1)$ on $[0, 1]$, and using Brent's MP package [4] to handle the multiple-precision computations on a VAX 11/780 in the Department of Mathematical Sciences at Kent State University, the iterates of the Remez algorithm were terminated when $\|\sqrt{t} - h_n(t)\|_{L_\infty[0,1]}$ and $|\lambda|$ agreed [cf. (2.9)] to 100 decimal digits. Because of quadratic convergence, at most *nine* iterations of the Remez algorithm were needed for convergence in each case considered. Taking into account guard digits and the possibility of some small rounding errors, we believe that the numbers $\{E_{2n}(|x|)\}_{n=1}^{52}$ we determined are accurate to at least 95 decimal places.

To conserve space, we give the products $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ to 20 decimal digits in Table 2.1 to show the slow convergence of this sequence. (Printouts of $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ to 100 decimal digits are available on request.)

It appears that the products $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ in Table 2.1 have converged to four decimal digits. Based on the asymptotic estimates of Section 5, in order to have $|2nE_{2n}(|x|) - \beta| < 10^{-6}$, one would need $n \geq 210$. This truly formidable computation would require finding polynomials of best uniform approximation to $\sqrt{t}$ on $[0, 1]$ of degree at least 210.

3. Computing Upper Bounds for the Bernstein Constant β

An ingenious step in Bernstein's analysis [2] of the convergence [cf. (1.10)] of the sequence $\{2nE_{2n}(|x|)\}_{n=1}^{\infty}$ was the introduction of the real function $F(t)$, defined by

$$(3.1) \qquad F(t) := \sum_{k=0}^{\infty} \frac{t}{(t + 2k + 1)^2 - \frac{1}{4}} \qquad (t \geq 0).$$

As shown by Bernstein [2], this function admits a representation in terms of the psi (digamma) function $\psi(z)$ (cf. Whittaker and Watson [12, p. 240]),

$$(3.2) \qquad \psi(z) := \frac{d}{dz}(\log \Gamma(z)) = \frac{\Gamma'(z)}{\Gamma(z)},$$

by means of

$$(3.3) \qquad F(t) = \frac{t}{2}\left\{\psi\left(\frac{t}{2} + \frac{3}{4}\right) - \psi\left(\frac{t}{2} + \frac{1}{4}\right)\right\} \qquad (t \geq 0).$$

Other representations for $F(t)$ (cf. [2]) include

$$(3.4) \quad F(t) = \frac{t}{2t+1} F\left(1, 1; t + \frac{3}{2}; \frac{1}{2}\right) = t\int_0^1 \frac{z^{t-1/2}\,dz}{z+1} = \frac{1}{2}\int_0^{\infty} \frac{e^{-u}\,du}{\cosh(u/2t)},$$

where $F(a, b; c; z)$ denotes the classical hypergeometric function (cf. Henrici [7, p. 27]). The last integral in (3.4) shows that $F(t)$ is strictly increasing on $[0, +\infty)$, with $F(0) = 0$ and $F(+\infty) = \frac{1}{2}$.

The connection between the function $F(t)$ and the Bernstein constant β of (1.10) is the following. For each positive integer m, set

(3.5)

$$\mu_m := \inf_{\substack{a_0,\ldots,a_m \\ \text{real}}} \left\{ \left\| \cos(\pi t) \left[F(t) - \left(a_0 + \sum_{k=1}^{m} \frac{a_k}{t^2 - [(2k-1)/2]^2} \right) \right] \right\|_{L_\infty[0,+\infty)} \right\},$$

and for $m = 0$, set

(3.5′) $$\mu_0 := \inf_{a_0 \text{ real}} \left\{ \left\| \cos(\pi t)[F(t) - a_0] \right\|_{L_\infty[0,+\infty)} \right\}.$$

Note that

(3.6) $$\lim_{t \to k-1/2} \frac{\cos(\pi t)}{t^2 - [(2k-1)/2]^2} = \frac{(-1)^k \pi}{2k-1} \qquad \text{(all positive integers } k\text{)},$$

so that the poles of the sum in (3.5) are cancelled by zeros of $\cos(\pi t)$. As a consequence, by standard arguments there exist real constants $\{\hat{a}_k(m)\}_{k=0}^{m}$ such that

(3.7) $$\mu_m = \left\| \cos(\pi t) \left[F(t) - \left(\hat{a}_0(m) + \sum_{k=1}^{m} \frac{\hat{a}_k(m)}{t^2 - [(2k-1)/2]^2} \right) \right] \right\|_{L_\infty[0,+\infty)}$$

It is evident from the definition of (3.5) that the sequence of positive constants $\{\mu_m\}_{m=0}^{\infty}$ is nonincreasing:

(3.8) $$\mu_0 \geq \mu_1 \geq \mu_2 \geq \cdots .$$

Now, Bernstein [2, p. 55] proved that the Bernstein constant β of (1.10) and the limit of the sequence $\{\mu_m\}_{m=0}^{\infty}$ are connected through

(3.9) $$\beta = 2 \lim_{m \to \infty} \mu_m.$$

Thus, with (3.8), each constant μ_m of the approximation problem (3.5) provides an upper bound for β:

(3.10) $$\beta \leq 2\mu_m \qquad \text{(each nonnegative integer } m\text{)}.$$

Because of the monotone character of $F(t)$ on $[0, +\infty)$ and the fact that $F(0) = 0$ and $F(+\infty) = \frac{1}{2}$, it immediately follows that $\mu_0 = \frac{1}{4}$, whence

(3.11) $$2\mu_0 = \tfrac{1}{2}.$$

Bojanic and Elkins [3] numerically estimated the solution of (3.5) for $m = 1$ and claimed that

(3.12) $$2\mu_1 \leq 0.30981\,60614.$$

Interestingly, Bernstein [2] numerically estimated the solution of (3.5) for $m = 3$ in 1914 and found that $\mu_3 < 0.143$, so that

(3.13) $$2\mu_3 < 0.286,$$

which is the upper bound for β mentioned in (1.11). (More accurate estimates of $2\mu_1$ and $2\mu_3$ can be found in Table 3.1.)

It turns out that the solution of the approximation problem in (3.5) has an interesting oscillation character that permits the use of a modified form of the (second) Remez algorithm. (It should be mentioned that the work of Bernstein [2] in 1914 *predates* the 1934 appearance of Remez's algorithm [9]!) The minimization problem (3.5) was solved using the following modified (second) Remez algorithm:

Step 1. Let $\tilde{S} := \{t_j\}_{j=0}^{m+1}$ $(m \geq 1)$, be a set of $m+2$ distinct (alternation) points in $[0, +\infty]$ satisfying

$$0 =: t_0 < t_1 < \ldots < t_m \leq m - \tfrac{1}{2} < t_{m+1} := +\infty. \tag{3.14}$$

Step 2. Find the $m+2$ unique constants $\{a_i\}_{i=0}^{m}$ and λ, a linear problem, such that

(3.15)

$$\begin{cases} \cos(\pi t_j)\left\{a_0 + \sum_{k=1}^{m} \dfrac{a_k}{t_j^2 - [(2k-1)/2]^2}\right\} + (-1)^{j+1}\lambda = \cos(\pi t_j) F(t_j) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (j = 0, \ldots, m), \\ a_0 + \lambda = \tfrac{1}{2} = F(+\infty). \end{cases}$$

The solution of the linear problem (3.15) forces the function

$$R_m(t) := \cos(\pi t)\left[F(t) - \left(a_0 + \sum_{k=1}^{m} \frac{a_k}{t^2 - [(2k-1)/2]^2}\right)\right] \tag{3.16}$$

to equioscillate on the subset $\{t_j\}_{j=0}^{m}$ of $\tilde{S}$ with (alternating) error $|\lambda|$, and, in addition, forces $R_m(t)$ to oscillate between $+|\lambda|$ and $-|\lambda|$ as $t \to \infty$ (see Fig. 3.1). Thus, in analogy with (2.8),

$$\|R_m(t)\|_{L_\infty(\tilde{S})} = |\lambda|, \tag{3.17}$$

and because $\tilde{S}$ is a subset of $[0, +\infty]$, then [cf. (2.9)]

$$\|R_m(t)\|_{L_\infty[0,+\infty)} - |\lambda| \geq 0. \tag{3.18}$$

As background for the conditions of the next step of this modified Remez algorithm, we remark that

$$\begin{aligned} &\frac{d}{dt}\left\{F(t) - \left(a_0 + \sum_{k=1}^{m} \frac{a_k}{t^2 - [(2k-1)/2]^2}\right)\right\} \\ &\qquad = F'(t) + 2t \sum_{k=1}^{m} \frac{a_k}{\left[t^2 - [(2k-1)/2]^2\right]^2}. \end{aligned} \tag{3.19}$$

Table 3.1. $\{2\mu_m\}$.

m	$2\mu_m$	m	$2\mu_m$
0	0.50000 00000 00000	10	0.28056 81480 84662
1	0.30981 66482 77486	20	0.28026 79181 28026
2	0.28964 46428 36759	30	0.28021 30013 47551
3	0.28458 56232 64382	40	0.28019 38951 81171
4	0.28268 16444 08752	50	0.28018 50827 23738
5	0.28177 99926 24272	60	0.28018 03067 66681
6	0.28128 65208 69723	70	0.28017 74317 42434
7	0.28098 84334 65837	80	0.28017 55680 33390
8	0.28079 50582 78019	90	0.28017 42915 00582
9	0.28066 26720 87176	100	0.28017 33791 01718

Thus, if the solution of (3.15) is such that the a_j's $(0 \le i \le m)$ and λ are all positive numbers, then the monotone character of $F(t)$ gives, with (3.19), that

$$G_m(t) := F(t) - \left(a_0 + \sum_{k=1}^{m} \frac{a_k}{t^2 - [(2k-1)/2]^2}\right)$$

is strictly increasing from $-\infty$ to $+\lambda$ on the interval $(m - \frac{1}{2}, +\infty)$. Defining $\tau_m = \tau_m(\tilde{S})$ to be the unique value of t in $(m - \frac{1}{2}, +\infty)$ such that $G_m(\tau_m) = -\lambda$, then because $|R_m(t)| = |\cos(\pi t) \cdot G_m(t)| \le |G_m(t)| \le \lambda$ for all $t \ge \tau_m$, and because $G_m(t) \to \frac{1}{2} - a_0 = \lambda$ as $t \to +\infty$, it follows that

$$\|R_m(t)\|_{L_\infty[\tau_m, +\infty)} = \lambda.$$

On the other hand, since $\{t_j\}_{j=0}^{m}$ is a subset of $[0, m - \frac{1}{2}]$, then from (3.17) $\|R_m(t)\|_{L_\infty[0, m-1/2]} \ge \lambda$. Hence, since $[0, m - \frac{1}{2}]$ is a subset of $[0, \tau_m]$,

$$\|R_m(t)\|_{L_\infty[0, \tau_m(\tilde{S})]} = \|R_m(t)\|_{L_\infty[0, +\infty)}, \tag{3.20}$$

and it can then be numerically determined whether the stronger statement

$$\|R_m(t)\|_{L_\infty[0, m-1/2]} = \|R_m(t)\|_{L_\infty[0, +\infty)} \tag{3.20$'$}$$

is valid or not. [We remark that in the special cases treated thus far in the literature for determining upper bounds for β, such as $m = 1$ of Bojanic and Elkins [3] and $m = 3$ of Bernstein [2], the choices of the associated sets $\tilde{S}$ were such that the a_j's $(0 \le j \le m)$ and λ of step 2 were positive, and (3.20′) was also satisfied.] This brings us to

Step 3. If the solutions $\{a_i\}_{i=0}^{m}$ and λ of step 2 are all positive numbers, if (3.20′) is satisfied, and if $\|R_m(t)\|_{L_\infty[0, m-1/2]} - \lambda$ is sufficiently small, the iteration is terminated. Otherwise, find a new alternation set $\tilde{S}$ consisting of a set of m local extrema of $R_m(t)$ (with alternating signs) in $(0, m - \frac{1}{2}]$, in addition to $t_0 := 0$ and $t_{m+1} := \infty$, and repeat Steps 2 and 3, etc.

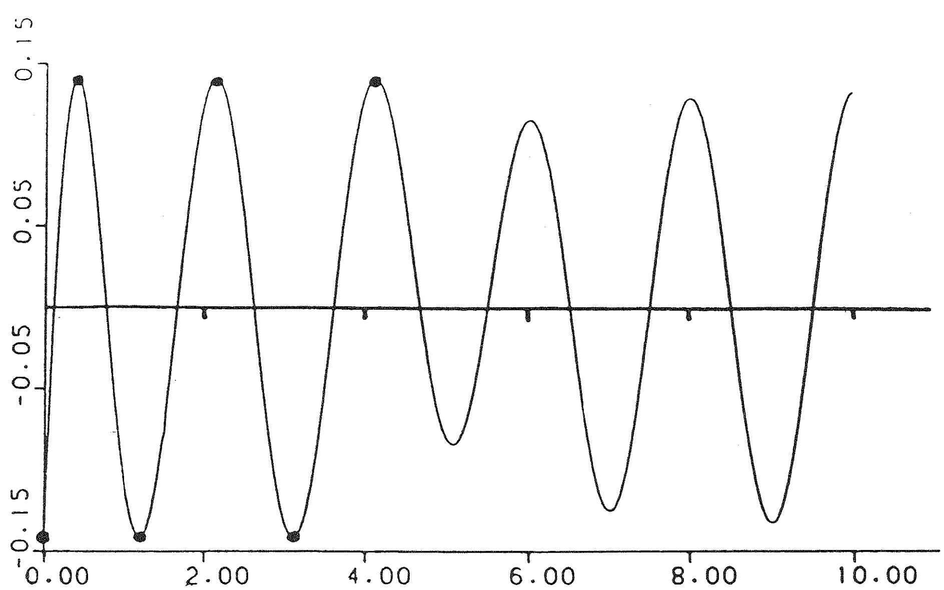

Fig. 3.1. $R_5(t)$.

Starting with the particular set $\tilde{S}^{(0)} := \{t_j^{(0)}\}_{j=0}^{m+1}$, where

$$(3.21)\quad t_0^{(0)} := 0; \qquad t_j^{(0)} := \frac{2j-1}{2} \qquad (j = 1, 2, \ldots, m); \qquad t_{m+1}^{(0)} := +\infty,$$

the associated system of linear equations (3.15) was solved for $\{a_j^{(0)}\}_{j=0}^{m}$ and $\lambda^{(0)}$, using Gaussian elimination with partial pivoting. In every case considered in the tabulation in Table 3.1, the starting values (3.21) for the set $\tilde{S}$ of alternation points were sufficiently good so that Steps 2 and 3 of the above modified Remez algorithm always produced positive numbers $\{a_k\}_{k=0}^{m}$ and λ, as well as alternation sets satisfying (3.14) and (3.20′). Moreover, the convergence of this algorithm was, as might be expected, quadratic, and at most ten iterations of this modified Remez algorithm were needed for convergence in the cases considered.

In Fig. 3.1, we graph the function $R_5(t)$ of (3.16), which is associated with the best approximation constant μ_5 of (3.7). In this figure, there are six alternation points (denoted by small dark disks) in the interval $[0, 9/2]$, as well as oscillations in $(9/2, +\infty)$ that grow in modulus to μ_5 as $t \to \infty$.

Unfortunately, the above calculations were *not* carried out to the same high accuracy (100 decimal digits) as were the calculations of $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ of Section 2. One reason for this is that the modified Remez algorithm applied to the minimization problem of (3.5) necessarily requires repeated evaluation of the function $F(t)$ of (3.1). Here, we used the representation (3.3) of $F(t)$ in terms of the psi function $\psi(x)$, and $\psi(x)$ was approximated, based on the work of Cody, Strecok, and Thacher [6], by

$$\psi(x) \doteq (x - x_0) R_{8,8}(x), \qquad \text{for } \tfrac{1}{2} \le x \le 3, \tag{3.22}$$

where $x_0 = 1.46163\ldots$ is the unique positive zero of $\psi(x)$, known to 40 significant digits, and where $R_{8,8}(x)$ is the ratio of two specific polynomials in x of degree 8 from [6]; and by

$$\psi(x) \doteq \ln x - \frac{1}{2x} + R_{6,6}\left(\frac{1}{x^2}\right) \qquad (\text{for } 3 \le x < \infty), \tag{3.23}$$

where $R_{6,6}(u)$ is the ratio of two specific polynomials in u of degree 6 from [6]. For the range $0 < x \le \frac{1}{2}$, we used the known recurrence relation for the function $\psi(x)$

$$\psi(x) = \psi(x+1) - \frac{1}{x}, \tag{3.24}$$

and the approximation of (3.22).

Now, the above approximations of $\psi(t)$ are good to about 20 significant digits (cf. [6]), so we estimate that our calculations of $\{2\mu_m\}_{m=0}^{100}$ are accurate to at least 15 decimal digits. To conserve space, a subset of the numbers $\{2\mu_m\}_{m=0}^{100}$ is given to 15 decimals in Table 3.1.

It is evident from Table 3.1 that from (1.12) and (3.8)–(3.9),

$$\frac{1}{2\sqrt{\pi}} = 0.28209\,47917\cdots > 2\mu_5 > \beta, \tag{3.25}$$

so that the Bernstein Conjecture is *false*!

4. Computing Lower Bounds for the Bernstein Constant β

We have seen that our calculations in Section 3 of the upper bounds $\{2\mu_m\}_{m=0}^{100}$ of the Bernstein constant β [cf. (1.10)] are *sufficient* to disprove Bernstein's conjecture (1.13). Thus, in terms of settling Bernstein's conjecture, it is obviously unnecessary to determine lower bounds for β. However, for completeness, we include here calculations of lower bounds for β based on another ingenious method of Bernstein. The calculations of these lower bounds for β proved to be the most laborious of all the calculations we performed.

To describe Bernstein's method [2] of determining lower bounds for β, we define

$$\phi_m(x) := \prod_{j=1}^{m-1} (x^2 - j^2) \qquad (m \ge 1) \tag{4.1}$$

and

$$\psi_m(x) = \psi_m(x; \lambda_1, \lambda_2, \ldots, \lambda_m) := \prod_{j=1}^{m} \left(x^2 - \lambda_j^2\right) \qquad (m \ge 1). \tag{4.2}$$

(Here, we use the convention that $\prod_{j=\beta}^{\alpha} := 1$ if $\alpha < \beta$.) The parameters appearing in (4.2) are assumed to satisfy

$$j - 1 < \lambda_j < j \qquad (j \ge 1). \tag{4.3}$$

Then, for each $m \geq 1$, set

$$(4.4)\quad B_m(\lambda_1, \lambda_2, \ldots, \lambda_m) := \frac{\displaystyle\sum_{i=1}^{m} \frac{\phi_m(\lambda_i)}{\psi_m'(\lambda_i)}\left[1 - \left(\frac{2\lambda_i}{\lambda_i + \frac{1}{2}}\right) F\left(\lambda_i + \frac{1}{2}\right)\right]}{\displaystyle\sum_{i=1}^{m} \frac{\phi_m(\lambda_i)}{\psi_m'(\lambda_i)}\left[\frac{2}{\pi\lambda_i} + \tan\left(\frac{\pi}{2}[\lambda_i - i + 1]\right)\right]},$$

where the function $F(t)$ is defined in (3.1). Note that from (4.1) and (4.2), we can write

$$(4.5)\quad \frac{\phi_m(\lambda_i)}{\psi_m'(\lambda_i)} = \frac{\prod_{j=1}^{i-1}(\lambda_i^2 - j^2) \cdot \prod_{j=i}^{m-1}(j^2 - \lambda_i^2)}{2\lambda_i \prod_{j=1}^{i-1}(\lambda_i^2 - \lambda_j^2) \cdot \prod_{j=i+1}^{m}(\lambda_j^2 - \lambda_i^2)} \qquad (1 \leq i \leq m).$$

With the conditions of (4.3), we see that the above ratios are all positive. Thus, since $F(t)$ is monotone increasing on $[0, +\infty)$ with $F(0) = 0$ and with $F(+\infty) = \frac{1}{2}$ (cf. Section 3), we deduce that each term of either sum of (4.4) is necessarily positive, whence $B_m(\lambda_1, \lambda_2, \ldots, \lambda_m) > 0$.

With β as defined in (1.10), Bernstein [2] showed that $B_m(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is a lower bound for β; i.e.,

$$(4.6)\quad \beta \geq B_m(\lambda_1, \lambda_2, \ldots, \lambda_m),$$

for each positive integer m and for each choice of $\{\lambda_j\}_{j=1}^{m}$ satisfying (4.3). The best such lower bound for β for each $m \geq 1$ is evidently given by

$$(4.7)\quad l_m := \sup\left\{B_m(\lambda_1, \lambda_2, \ldots, \lambda_m) : \{\lambda_j\}_{j=1}^{m} \text{ satisfies (4.3)}\right\},$$

so that

$$(4.8)\quad \beta \geq l_m > 0 \qquad (m \geq 1).$$

Next, consider any parameters $\{\lambda_j\}_{j=1}^{m+1}$ satisfying (4.3). On fixing $\{\lambda_j\}_{j=1}^{m}$ and on letting λ_{m+1} decrease to m, it can readily be verified from (4.4) and (4.5) that

$$(4.9)\quad \lim_{\lambda_{m+1} \to m} B_{m+1}(\lambda_1, \lambda_2, \ldots, \lambda_m, \lambda_{m+1}) = B_m(\lambda_1, \lambda_2, \ldots, \lambda_m).$$

As a consequence, we see from (4.7) that

$$(4.10)\quad l_{m+1} \geq l_m \qquad (m \geq 1),$$

so that, with (4.8), $\{l_m\}_{m=1}^{\infty}$ is a bounded nondecreasing sequence of positive numbers. Now, Bernstein [2] further showed that the limit of this sequence is β:

$$(4.11)\quad \beta = \lim_{m \to \infty} l_m.$$

Bernstein in fact numerically estimated l_1 and l_2 and found that (cf. [2])

$$(4.12)\quad l_1 > 0.27 \qquad \text{and} \qquad l_2 > 0.278.$$

This last estimate of l_2 appears as the lower bound of β in (1.11). (More accurate estimates of l_1 and l_2 can be found in Table 4.1.)

Table 4.1. $\{l_m\}_{m=1}^{20}$.

m	l_m	m	l_m
1	0.27198 23590 30477	11	0.28016 34641 87524
2	0.27893 09228 49406	12	0.28016 48933 27009
3	0.27981 10004 37231	13	0.28016 59052 38063
4	0.28002 43339 28903	14	0.28016 66415 27680
5	0.28009 77913 15214	15	0.28016 71898 92928
6	0.28012 91830 79687	16	0.28016 76066 00825
7	0.28014 46910 09336	17	0.28016 79288 71653
8	0.28015 31877 11753	18	0.28016 81819 90114
9	0.28015 82176 99044	19	0.28016 83835 39180
10	0.28016 13794 71687	20	0.28016 85460 02042

We now describe our calculations of the lower bounds l_m. It is evident from (4.4) that the parameters $\lambda_1, \lambda_2, \ldots, \lambda_m$ enter nonlinearly in the definition of $B_m(\lambda_1, \lambda_2, \ldots, \lambda_m)$. Computationally, we used a fairly standard optimization (maximization) routine, without derivatives, to optimize the real parameters $\{\lambda_i\}_{i=1}^m$, subject to the constraints of (4.3), to determine l_m. Again, because the function $F(t)$ appears explicitly in the definition of $B_m(\lambda_1, \lambda_2, \ldots, \lambda_m)$ of (4.4), we used the approximations of (3.22)–(3.24) for the psi function $\psi(t)$ and the representation (3.3) of $F(t)$ in terms of $\psi(t)$. As in our calculations of the upper bounds $2\mu_m$ of β [cf. (3.10)], our calculations of the lower bounds l_m of β were *not* carried out to the same high accuracy as were the calculations of $2nE_{2n}(|x|)$ (100 decimal digits) in Section 2. For reasons similar to those applying to the numerical results of Section 3, since our approximations of $\psi(t)$ are good to about 20 significant digits (cf. [6]), we estimate that the optimization calculations of $\{l_m\}_{m=1}^{20}$ are accurate to at least 15 decimal digits. The numbers $\{l_m\}_{m=1}^{20}$ are given to 15 decimal places in Table 4.1.

On comparing the upper bounds of Table 3.1 with the lower bounds of Table 4.1, we see that the lower bound l_m of (4.7) is a considerably more *accurate* estimate of β of (1.18) than is the upper bound $2\mu_m$ of (3.5), for *each* $1 \leq m \leq 20$. In fact, the error in l_{20} in approximating β is roughly $9.53 \cdot 10^{-7}$, while that of $2\mu_{100}$ is only $3.88 \cdot 10^{-6}$. However, this gain in accuracy was largely offset by the increased computer time necessary to find the numbers l_m by our optimization routine. We also remark that this greater accuracy of the lower bounds l_m in approximating β explains why the m values in Table 4.1 do not range as high as those for the upper bounds $2\mu_m$ in Table 3.1.

5. The Richardson Extrapolation of the Numbers $\{2nE_{2n}(|x|)\}_{n=1}^{52}$

The numbers $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ appearing in Table 2.1 indicate that the convergence [cf. (1.10)] of these numbers to the Bernstein constant β is quite slow. One typical scheme for improving the convergence rate of slowly convergent sequences is the *Richardson extrapolation method* (cf. Brezinski [5, p. 7]), which can be described as follows. If $\{S_n\}_{n=1}^N$, where $N > 2$, is a given (finite) sequence of real numbers, set $T_0^{(n)} := S_n$ $(1 \le n \le N)$, and regard $\{T_0^{(n)}\}_{n=1}^N$ as the zero-th column, consisting of N numbers, of the Richardson extrapolation table. The first column of the Richardson extrapolation table, consisting of $N - 1$ numbers, is defined by

$$(5.1) \qquad T_1^{(n)} := \frac{x_n T_0^{(n+1)} - x_{n+1} T_0^{(n)}}{x_n - x_{n+1}} \qquad (1 \le n \le N-1),$$

and inductively, the $(k+1)$-st column of the Richardson extrapolation table, consisting of $N - k - 1$ numbers, is defined by

$$(5.2) \qquad T_{k+1}^{(n)} := \frac{x_n T_k^{(n+1)} - x_{n+k+1} T_k^{(n)}}{x_n - x_{n+k+1}} \qquad (1 \le n \le N-k-1),$$

for each $k = 0, 1, \ldots, N-2$, where $\{x_n\}_{n=1}^N$ are given constants. In this way, a triangular table, consisting of $N(N+1)/2$ entries, is created. In our case of $\{2nE_{2n}(|x|)\}_{n=1}^{52}$, a triangular table of 1,378 entries was created. As for the numbers $\{x_n\}_{n=1}^{52}$ of (5.1)–(5.2), preliminary calculations indicated that $2nE_{2n}(|x|) \doteqdot \beta + K/n^2 +$ lower-order terms, so we chose $x_n := 1/n^2$. We remark that the potential loss of accuracy in the subtractions in the numerators and denominators of the fractions defined in (5.1) and (5.2) suggested that the calculations of $2nE_{2n}(|x|)$ be done to very high precision (100 decimal digits).

The Richardson extrapolation of $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ produced unexpectedly beautiful results. Rather than presenting here the complete extrapolation table of 1,378 entries (giving each entry to, say, 95 decimal digits), it seems sufficient to mention that of the last 20 columns of this table, all but 3 of the 210 entries in these columns agreed with the first 45 digits of the following approximation of β in (1.18):

$$(5.3) \qquad \beta \doteqdot 0.28016\,94990\,23869\,13303\,64364\,91230\,67200\,00424\,82139\,81236.$$

In addition, 182 entries of the 210 entries in these 20 last columns of this Richardson extrapolation table agreed with *all 50* digits of the above approximation of β!

The success of this Richardson extrapolation (with $x_n := 1/n^2$) applied to $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ strongly suggests that $2nE_{2n}(|x|)$ admits an asymptotic expansion of the form

$$(5.4) \qquad 2nE_{2n}(|x|) \stackrel{?}{=} \beta - \frac{K_1}{n^2} + \frac{K_2}{n^4} - \frac{K_3}{n^6} + \ldots, \qquad n \to \infty,$$

where the constants K_j are independent of n. (This is the basis for the new conjecture of Section 1.) Assuming that (5.4) is valid, it follows that

$$(5.5) \qquad n^2(2nE_{2n}(|x|) - \beta) = -K_1 + \frac{K_2}{n^2} - \frac{K_3}{n^4} + \ldots, \qquad n \to \infty.$$

Thus, with the known high-precision approximations of $2nE_{2n}(|x|)$ of Table 2.1, and with an estimate for β determined from the last entry of the Richardson extrapolation table for $\{2nE_{2n}(|x|)\}_{n=1}^{52}$, we can again apply Richardson extrapo-

Table 5.1. $\{K_j\}_{j=1}^{10}$

j	K_j
1	0.04396 75288 8
2	0.02640 71687 7
3	0.03125 34264 6
4	0.05889 00165 7
5	0.16010 69971
6	0.59543 53151
7	2.92591 5470
8	18.49414 033
9	146.94301 23
10	1438.03271 7

lation to $\{n^2(2nE_{2n}(|x|) - \beta)\}_{n=1}^{52}$ (with $x_n = 1/n^2$) to obtain an extrapolated estimate for K_1 of (5.5). This bootstrapping procedure can be continued to give, via Richardson extrapolation, estimates for the successive K_j of (5.4). As might be suspected, there is a progressive loss of numerical accuracy in the successive determination of the K_j in this way.

In Table 5.1, we give estimates of $\{K_j\}_{j=1}^{10}$ rounded to ten significant digits. As Table 5.1 indicates, the constants K_j begin to grow quite rapidly. That these constants all turned out to be *positive* has been incorporated into the new conjecture of Section 1.

Finally, because the Bernstein constant β has a connection [cf. (3.9)] with particular rational approximations to the function $F(t)$, where $F(t)$ can be expressed [cf. (3.4)] in terms of the classical hypergeometric function, it is not implausible to believe that β, as well as the constants K_j of (5.4), might admit a closed-form expression in terms of the classical hypergeometric function and/or known mathematical constants.

Acknowledgments. We wish to thank sincerely Professor R. Bojanic (Ohio State University) for having interested us in Bernstein's conjecture and for his considerable assistance in citing to us the relevant literature on this problem. This research was supported by the Department of Energy.

References

1. R. A. Bell, S. M. Shah (1969): *Oscillating polynomials and approximations to* $|x|$. Publ. of the Ramanujan Inst. **1**:167–177.
2. S. Bernstein (1914): *Sur la meilleure approximation de* $|x|$ *par des polynomes de degrés donnés*. Acta. Math., **37**:1–57.
3. R. Bojanic, J. M. Elkins (1975): *Bernstein's constant and best approximation on* $[0, \infty)$. Publ. de l'Inst. Math., Nouvelle série, **18** (32):19–30.
4. Richard P. Brent (1978): *A FORTRAN multiple-precision arithmetic package*. Assoc. Comput. Mach. Trans. Math. Software, **4**:57–70.
5. C. Brezinski (1978): Algorithmes d'Accélération de la Convergence. Paris: Éditions Technip.

6. W. J. Cody, A. J. Strecok, H. C. Thacher, Jr. (1973): *Chebyshev approximations for the psi function*. Math. of Comp., **27**:123–127.
7. P. Henrici (1974): Applied and Computational Complex Analysis, vol. 1. New York: John Wiley & Sons.
8. G. Meinardus (1967): Approximations of Functions: Theory and Numerical Methods. New York: Springer-Verlag.
9. E. Ya. Remez (1934): *Sur le calcul effectiv des polynômes d'approximation de Tchebichef*. C.R. Acad. Sci. Paris, **199**:337–340.
10. T. J. Rivlin (1969): An Introduction to the Approximation of Functions. Waltham, Massachusetts: Blaisdell Publishing Co.
11. D. A. Salvati (1980): Numerical Computation of Polynomials of Best Uniform Approximation to the Function $|x|$. Master's Thesis, Ohio State University, 39 pp. Columbus, Ohio.
12. E. T. Whittaker, G. N. Watson (1962): A Course of Modern Analysis, 4th ed. Cambridge: Cambridge University Press.

R. S. Varga
Institute for Computational Mathematics
Kent State University
Kent, Ohio 44242

A. J. Carpenter*
Institute for Computational Mathematics
Kent State University
Kent, Ohio 44242

**Present address:* Department of Mathematical Sciences, Butler University, Indianapolis, Indiana 46208