Reprinted from:
NUMERICAL SOLUTION OF PARTIAL
DIFFERENTIAL EQUATIONS – II

Academic Press, Inc., New York and London

SOME RESULTS IN APPROXIMATION THEORY WITH APPLICATIONS TO NUMERICAL ANALYSIS*

Richard S. Varga**

1. Introduction

The object of this paper is to present results in two rather different areas of approximation theory, and to sketch their applications to numerical analysis. In Sections 2-4, we discuss results concerning Chebyshev rational approximations of reciprocals of certain entire functions (such as $f(z) = e^z$) on $[0, +\infty)$, and we show how these approximations can be used numerically in the solution of semi-discrete parabolic partial difference equations. We also discuss in Section 4 results of numerical experiments testing such Chebyshev semi-discrete approximations.

In Section 5, we discuss improved error bounds for spline and L-spline interpolation. These improved error bounds for spline interpolation are then used to

*This research was supported in part by AEC Grant AT(11-1)-2075.

**Kent State University.

deduce improved error bounds for Galerkin approximations of solutions of particular two-point nonlinear boundary value problems in Section 6.

2. Chebyshev Semi-Discrete Approximations of Parabolic Partial Difference Equations

As described in Cody, Meinardus, and Varga [3], consider any linear system of N coupled ordinary differential equations of the form

$$(2.1)\qquad \begin{cases} \dfrac{dc(t)}{dt} = -Ac(t) + g, & \forall t > 0, \\ c(0) = \tilde{c}, & \end{cases}$$

where A is assumed to be an $N \times N$ time-independent Hermitian positive definite matrix, and $c(t)$, g, and $\tilde{c}$ are N-vectors. Typically, such coupled equations can arise from semi-discrete approximations to linear parabolic partial differential equations, in which all spatial variables are differenced, but the time variable, t, is left continuous (cf. [14, Chapter 8]). The solution $c(t)$ of (2.1) is given explicitly by

$$(2.2)\qquad c(t) = A^{-1}g + \exp(-tA)\{\tilde{c} - A^{-1}g\}, \quad \forall t \geq 0,$$

where as usual $\exp(-tA) \equiv \sum_{k=0}^{\infty} (-tA)^k/k!$.

To define the Chebeyshev semi-discrete approximations of (2.1), we turn to the following approximation problem. If π_m denotes all real polynomials $p(n)$ of degree at most m, and $\pi_{m,n}$ analogously denotes all real rational functions $r_{m,n}(x) = p(x)/q(x)$ with $p \in \pi_m$, $q \in \pi_n$, then let

$$(2.3) \qquad \lambda_{m,n} \equiv \inf_{\pi_{m,n}} \| e^{-x} - r_{m,n}(x) \|_{L_\infty[0,+\infty]}$$

denote the minimum error in approximating e^{-x} on $[0, +\infty)$ in the uniform norm over $\pi_{m,n}$. These constants $\lambda_{m,n}$ are called the Chebyshev constants for e^{-x} with respect to $[0, +\infty)$. It is obvious that $\lambda_{m,n}$ is finite if and only if $0 \leq m \leq n$, and moreover, given any pair (m,n) of nonnegative integers with $0 \leq m \leq n$, it is known (cf. Meinardus [6]) that there exists a unique $\hat{r}_{m,n}(x) \in \pi_{m,n}$ (after dividing out possible common factors) with

$$(2.4) \qquad \hat{r}_{m,n}(x) \equiv \hat{p}_{m,n}(x)/\hat{q}_{m,n}(x) ,$$

and with $\hat{q}_{m,n}(x) > 0$ on $[0, +\infty)$, such that

$$(2.5) \qquad \lambda_{m,n} = \| e^{-x} - \hat{r}_{m,n}(x) \|_{L_\infty[0,+\infty)} .$$

Since $\hat{q}_{m,n}(tA)$ is a polynomial in the matrix A, it is evident that $\hat{q}_{m,n}(tA)$ is a Hermitian positive definite $N \times N$ matrix for any finite $t \geq 0$. Thus, we can define the $(m,n)^{th}$ Chebyshev approximation $c_{m,n}(t)$ of $c(t)$ of (2.2) as

$$(2.6) \qquad c_{m,n}(t) = A^{-1}g + \hat{r}_{m,n}(tA)\{\tilde{c} - A^{-1}g\} , \quad \forall t \geq 0 ,$$

$$= A^{-1}g + (\hat{q}_{m,n}(tA))^{-1}[\hat{p}_{m,n}(tA)\{\tilde{c} - A^{-1}g\}], \quad \forall t \geq 0 .$$

In other words, $c_{m,n}(t)$ is defined for each finite $t \geq 0$ as the solution v of the system of linear equations

$$(2.7) \qquad \hat{q}_{m,n}(tA) \cdot v = k(t) \equiv \hat{q}_{m,n}(tA) \cdot A^{-1}g + \hat{p}_{m,n}(tA)\{\tilde{c} - A^{-1}g\} .$$

To estimate the error in $c(t) - c_{m,n}(t)$, we use ℓ_2-vector norms, i.e., $\|v\|_2 = (v^*v)^{\frac{1}{2}}$. If $\{\lambda_i\}_{i=1}^N$ denote the (positive) eigenvalues of A, then the Hermitian character of A gives us for any $t \geq 0$ that

$$\|\exp(-tA) - \hat{r}_{m,n}(tA)\|_2 = \max_{1 \leq i \leq N} \left| e^{-t\lambda_i} - \hat{r}_{m,n}(t\lambda_i) \right| ,$$

where $\|\cdot\|_2$ denotes the induced operator norm (or spectral norm) relative to the ℓ_2-vector norm. But as $t\lambda_i \geq 0$ for all $1 \leq i \leq N$, it follows from the definition of $\lambda_{m,n}$ in (2.3) that

$$\begin{aligned} &\|c(t) - c_{m,n}(t)\|_2 \\ (2.8) \qquad &\leq \|\exp(-tA) - \hat{r}_{m,n}(tA)\|_2 \cdot \|\tilde{c} - A^{-1}g\|_2 \\ &\leq \lambda_{m,n} \|\tilde{c} - A^{-1}g\|_2 \qquad \forall t \geq 0 . \end{aligned}$$

In general, the error of the spatial discretization leading to (2.1) must also be bounded to give the total error (i.e., space and time) of the Chebyshev semi-discrete approximations. This is discussed, for example, in [3] in a particular case.

Unlike usual methods of time-discretization, such as Crank-Nicolson, which depended upon repeatedly taking small time steps $\Delta t = T/M$ to achieve precision at a time T, the Chebyshev semi-discrete approximation directly (in one step) gives an approximation at time T by simply setting $t = T$ in (2.5). The accuracy of this method is clearly dependent from (2.8) on how the Chebyshev constants $\lambda_{m,n}$ behave as $n \to \infty$. First, from (2.3), it is evident that

$$(2.9) \qquad 0 < \lambda_{n,n} \leqslant \lambda_{n-1,n} \leqslant \cdots \leqslant \lambda_{o,n}, \qquad \forall n \geqslant 0 .$$

In [3], it was shown in particular that

$$(2.10) \qquad \lambda_{o,n} \leqslant (2e^{\alpha})^{-n} \qquad \forall n \geqslant 0 ,$$

where $\alpha = 0.13923\cdots$ is the real solution of $2\alpha e^{2\alpha+1} = 1$. Thus, (2.10) shows us that the Chebyshev constants $\lambda_{o,n}$ converge <u>geometrically to zero</u> as $n \to \infty$. In [3], it was also shown that this convergence is <u>not</u> faster than geometric, in that

$$\overline{\lim_{n\to\infty}} (\lambda_{o,n})^{1/n} \geqslant \frac{1}{6} .$$

Because of (2.9) and (2.10), we can state these results in the following form.

<u>Theorem 1</u>. Let $\{m(n)\}_{n=o}^{\infty}$ be any sequence of nonnegative integers with $0 \leqslant m(n) \leqslant n$ for each $n \geqslant 0$. Then,

$$(2.11) \qquad \overline{\lim_{n\to\infty}} (\lambda_{m(n),n})^{1/n} \leq \frac{e^{-\alpha}}{2} = 0.43501\cdots ,$$

and

$$(2.12) \qquad \overline{\lim_{n\to\infty}} (\lambda_{o,n})^{1/n} \geq \frac{1}{6} .$$

3. Theoretical Extensions.

It is natural to ask if the geometric convergence to zero of the Chebyshev constants $\lambda_{m,n}$ for $\frac{1}{e^x}$ in (2.11) and (2.12) hold for a wider class of entire functions than just $f(x) = e^x$. A generalization of these results has recently been given by Meinardus and Varga [8], and can be described as follows.

Let $f(z) = \sum_{k=o} a_k z^k$ be an entire function with $M_f(r) \equiv \sup_{|z|\leq r} |f(z)|$ as its maximum modulus function. Then, f is said to be of perfectly regular growth (ρ,B) (cf. Boas [2, p. 8], Valiron [13, p. 45]) if and only if there exist two (finite) positive numbers ρ and B such that

$$(3.1) \qquad \lim_{r\to\infty} \frac{\ln M_f(r)}{r^\rho} = B .$$

We then state (cf. [4])

Theorem 2. Let $f(z) = \sum_{k=o}^{\infty} a_k z^k$ be an entire function of perfectly regular growth (ρ,B) with $a_k \geq 0$ $\forall k \geq 0$, and for any pair (m,n) of nonnegative integers with $0 \leq m \leq n$, let

(3.2) $$\lambda_{m,n} \equiv \inf_{\pi_{m,n}} \left\| \frac{1}{f(x)} - r_{m,n}(x) \right\|_{L_\infty[0,+\infty]}$$

be its associated Chebyshev constants. Then, for any sequence $\{m(n)\}_{n=0}^{\infty}$ of nonnegative integers with $0 \leqslant m(n) \leqslant n$ for each $n \geqslant 0$,

(3.3) $$\overline{\lim_{n\to\infty}} (\lambda_{m(n),n})^{1/n} \leqslant 2^{-1/\rho} < 1 .$$

Moreover,

(3.4) $$\overline{\lim_{n\to\infty}} (\lambda_{0,n})^{1/n} \geqslant 2^{-2-1/\rho} .$$

Thus, Theorem 2 establishes the geometric convergence to zero of the Chebyshev constants for $1/f(x)$ on $[0,+\infty)$ for all entire functions of perfectly regular growth with nonnegative Taylor coefficients. As special cases of Theorem 2, we have of course $f(z) = e^z$, $f(z) = \sinh(z^n)$, and $f(z) = J_n(iz)$, where J_n denotes the n^{th} Bessel function of the first kind. Note that for $f(z) = e^z$, for which $\rho = B = 1$ from (3.1), the results of (3.3)-(3.4) of Theorem 2, are slightly weaker than those of (2.11)-(2.12) of Theorem 1.

It should be mentioned that the proofs of Theorems 1 and 2 depend upon estimating

$$\frac{1}{s_n(x)} - \frac{1}{f(x)}$$

where $s_n(z) = \sum_{k=0}^{n} a_k z^k$ is the n^{th} partial sum of $f(z)$.

It is shown in fact in [8] that, under the hypotheses of Theorem 2,

$$\lim_{n\to\infty} \left(\left\| \frac{1}{s_n} - \frac{1}{f} \right\|_{L_\infty[0,+\infty]} \right)^{1/n} = 2^{-1/\rho} ,$$

so that the upper bound of (3.3) cannot be improved using this specific technique.

Upon examining the conclusions of Theorem 2, we see that the bounds of (3.3)-(3.4) depend upon ρ, but not upon B, and this would suggest the possibility of extensions of Theorem 1 to entire functions which are not of perfectly regular growth. Such extensions have recently been considered in Meinardus, Taylor, Reddy and Varga [7], and we state below a representative result.

<u>Theorem 3</u>. Let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ be an entire function of order ρ, i.e., $\overline{\lim_{r\to\infty}} \frac{\ln\ln M_f(r)}{\ln r} \equiv \rho$, and assume that

$$(3.5) \qquad \overline{\lim_{r\to\infty}} \frac{\ln M_f(r)}{r^\rho} \equiv B, \quad \underline{\lim_{r\to\infty}} \frac{\ln M_f(r)}{r^\rho} \equiv b$$

satisfy $0 < b \leqslant B < \infty$, and that $a_k \geqslant 0$ $\forall k \geqslant 0$ with

$$(3.6) \qquad \frac{a_n}{a_{n+1}}$$

nondecreasing and unbounded for all n sufficiently large. Then, for any sequence $\{m(n)\}_{n=0}^{\infty}$ of nonnegative integers with $0 \leqslant m(n) \leqslant n$ for each $n \geqslant 0$, (3.7) $$\overline{\lim_{n\to\infty}} \left(\lambda_{m(n),n}\right)^{1/n} \leq \left(\frac{B}{2b}\right)^{1/\rho},$$

where the $\lambda_{m,n}$ are the Chebyshev constants of $1/f$, defined in (3.2). Moreover,

(3.8) $$\overline{\lim_{n\to\infty}} \left(\lambda_{o,n}\right)^{1/n} \geq \frac{1}{4}\left(\frac{b}{2B}\right)^{1/\rho}.$$

The results of Theorems 2 and 3 give then sufficient conditions on $f(z)$ so that its associated Chebyshev constants $\lambda_{m(n),n}$, $0 \leq m(n) \leq n$, converge geometrically to zero as $n \to \infty$. In the spirit of Bernstein's classical inverse-type theorems for polynomial and trigonometric polynomial approximation on finite intervals, the following result of [7] which we sketch, gives necessary conditions for this geometric convergence.

Theorem 4. Let $f(x) > 0$ be a real continuous function on $[0, +\infty)$, such that there exist a sequence of real polynomials $\{p_n(x)\}_{n=o}^{\infty}$ with $p_n \in \pi_n$ $\forall n \geq 0$ and a real number $q > 1$ such that

(3.9) $$\overline{\lim_{n\to\infty}} \left(\left\| \frac{1}{p_n(x)} - \frac{1}{f(x)} \right\|_{L_\infty[o,+\infty]} \right)^{1/n} = \frac{1}{q} < 1 .$$

Then, there exists a function $F(z)$ with $F(x) = f(x)$ $\forall x \geq 0$ such that F is analytic in the whole complex plane, i.e., F is entire. Moreover, F is of finite order, i.e.,

$$\overline{\lim_{r\to\infty}} \frac{\ln \ln M_F(r)}{\ln r} = \rho < \infty .$$

Proof. For any q_1 with $q > q_1 > 1$, it follows from (3.9) that there exists a positive integer $n_1(q_1)$ such that

$$\left\| \frac{1}{p_n} - \frac{1}{f} \right\|_{L_\infty[0,+\infty]} \leq \frac{1}{q_1^n}, \quad \forall n \geq n_1(q_1),$$

or equivalently,

$$(3.10) \qquad -\frac{1}{q_1^n} \leq \frac{1}{p_n(x)} - \frac{1}{f(x)} \leq \frac{1}{q_1^n}, \quad \forall n \geq n_1(q_1), \quad \forall x \geq 0.$$

Next, define

$$(3.11) \qquad m_f(r) = \|f\|_{L_\infty[0,r]}, \quad \text{where } 0 \leq r < +\infty.$$

Fixing $r > 0$, the fact that q_1 exceeds unity implies that there exists a positive integer $n_2(r)$ such that

$$(3.12) \qquad q_1^n \geq q_1^n - m_f(r) \geq \frac{q_1^n}{2}, \quad \forall n \geq n_2(r).$$

With $n_3 \equiv \max(n_1(q_1), n_2(r))$, a simple manipulation of (3.10) gives

$$\frac{-f^2(x)}{q_1^n + f(x)} \leq p_n(x) - f(x) \leq \frac{f^2(x)}{q_1^n - f(x)}, \quad \forall 0 \leq x \leq r, \quad \forall n \geq n_3.$$

From these inequalities and the inequalities of (3.12), it follows that

$$(3.13)\qquad \|p_n - f\|_{L_\infty[0,r]} \leq \frac{m_f^2(r)}{q_1^n - m_f(r)} \leq \frac{2m_f^2(r)}{q_1^n}\ ,\quad \forall n \geq n_3\ .$$

We now make the change of variables

$$\frac{r}{2}(1+t) = x\ ;\quad 0 \leq x \leq r\ ,\quad -1 \leq t \leq +1\ ,$$

and define

$$(3.14)\qquad h(t;r) = f\left(\frac{r}{2}(1+t)\right)\ .$$

If $E_n\{h(\cdot,r)\} \equiv \inf_{\sigma_n \in \pi_n} \|\sigma_n - h(\cdot,r)\|_{L_\infty[-1,+1]}$ denotes the error in the best Chebyshev polynomial approximation in π_n to $h(x,r)$ on $[-1,+1]$, the inequality of (3.13) immediately gives us that

$$(3.15)\qquad E_n\{h(\cdot,r)\} \leq \frac{2m_f^2(r)}{q_1^n}\ ,\quad \forall n \geq n_3\ .$$

Since $r > 0$ is fixed and q_1 is an arbitrary number with $q > q_1 > 1$, we evidently have from (3.15) that

$$(3.16)\qquad \overline{\lim_{n\to\infty}}\,(E_n\{h(\cdot,r)\})^{1/n} \leq \frac{1}{q}\ ,\quad \forall r > 0\ .$$

Using Bernstein's Theorem (cf. [6, p. 86]), it follows that $h(t;r)$ can, for each $r > 0$, be extended to a function analytic in the open ellipse $\mathcal{E}_q$ with foci at ± 1 and semi-major and semi-minor axes a and b such that

$a + b = q > 1$. In terms of f , this means that f can be extended to a function $F(z)$ analytic in the region $\Omega_r = \{z: z = \frac{r}{2}(1+t) \text{ where } t \in \mathcal{E}_q\}$. But, r is an arbitrary positive real number, and it is easily seen that for any complex number w , $w \in \Omega_r$ for r sufficiently large. Hence, $F(z)$ is analytic in the whole complex plane, i.e., F is an entire function, which proves the first assertion of Theorem 4. The proof that F has finite order, which depends on comparing $\sup_{\Omega_r} |F(z)|$ with $m_f(r)$ as $r \to \infty$, is only slightly more difficult, and can be found in [7].

Not all entire functions f which are real and positive on $[0,+\infty)$ satisfy the geometric convergence of (3.9). As shown in [7], the particular entire function

$$f(z) = (z + 1)\{2 + \cos z\}$$

which is real and positive on $[0,+\infty)$, fails to satisfy (3.9).

4. Numerical Results

Dr. W. E. Culham of the Gulf Research and Development Co. (Pittsburgh) has been numerically testing the Chebyshev semi-discrete method of Sections 2-3 for solving linear parabolic problems with one spatial variable. Though the results are not yet complete, some interesting conclusions can already be made.

First, for very small $T > 0$, it is generally preferable to use the Crank-Nicolson method rather than the Chebyshev semi-discrete method. The reason for this is almost obvious. The Crank-Nicolson method for one spatial variable problems can be viewed as giving a matrix Padé approximation $M(t)$ of $\exp(-tA)$ for which (cf. [14, p. 266])

$$\exp(-tA) = M(t) + O(t^3), \quad t \downarrow 0 .$$

In other words, $M(t)$ is a third-order approximation for $\exp(-tA)$ for t close to zero. The Chebyshev semi-discrete method, on the other hand, gives a matrix approximation of $\exp(-tA)$ for which the maximum error occurs at $t = 0$. This is a consequence of the Chebyshev equi-oscillation of the error curve. More precisely, we necessarily have (cf. (2.5)) that

$$\lambda_{m,n} = |e^{-\hat{x}} - \hat{r}_{m,n}(\hat{x})| \quad \text{for} \quad \hat{x} = 0 ,$$

and consequently,

$$\|\exp(-tA) - \hat{r}_{m,n}(tA)\|_2 = \lambda_{m,n} \quad \text{for} \quad t = 0 .$$

This short-coming of the Chebyshev semi-discrete method to small t can be partially off-set by using the following suggestion of Professor R. B. Kellogg. From the error curve $e^{-x} - \hat{r}_{m,n}(x)$, which necessarily has $m + n + 1$ distinct positive zeroes, let $\sigma_{m,n} > 0$ be the smallest such positive zero. Then, with $x = \sigma_{m,n} + t$, it follows from $|e^{-x} - \hat{r}_{m,n}(x)| \leq \lambda_{m,n}$, $\forall x \geq 0$ that

$$|e^{-t} - e^{\sigma_{m,n}} \hat{r}_{m,n}(\sigma_{m,n} + t)| \leq e^{\sigma_{m,n}} \lambda_{m,n}, \quad \forall t \geq 0,$$

and the rational approximation $e^{\sigma_{m,n}} \hat{r}_{m,n}(\sigma_{m,n}+t)$ of e^{-t} has zero error at $t=0$. This is equivalent with solving the Chebyshev minimization problem over $\pi_{m,n}$ for e^{-t} on $[0,+\infty)$ with the linear constraint of zero error at $t=0$. The numbers $\sigma_{n,n}$ have been computed by W. J. Cody, Jr., and they decrease very rapidly to zero as $n \to \infty$.

Because of the above-mentioned error behavior at $t=0$, the numerical experiments comparing the Crank-Nicolson method with the Chebyshev semi-discrete method have centered about comparing total work on a computer for T large. Physically speaking, T in these experiments is selected to be about the half-life of the transient term. As previously mentioned, though the results are not complete, several typical cases have arisen where the Chebyshev semi-discrete method with $m=n=3$ is about 100 times faster than the Crank-Nicolson method. More will be reported on this at a later time.

It should be stated that these Chebyshev semi-discrete methods as described are rather severely limited to linear problems for which the natural semi-group property for such parabolic problems holds. This is the essence of approximating e^{-x} on $[0,+\infty)$ by rational functionals in the uniform norm. It is not known if such techniques can be extended to the numerical solution of strongly nonlinear parabolic partial differential equations.

Finally, we wish to comment on the practical solution of the matrix equation of (2.7). Because $\hat{q}_{m,n}(x)$ is positive on $[0,+\infty)$, it can be factored in the form

$$\hat{q}_{m,n}(x) = \prod_{\ell=1}^{k_1} j_\ell(x) \cdot \prod_{\ell=1}^{k_2} h_\ell(x) , \quad 2k_2+k_1 = n ,$$

where each $j_\ell(x)$ is real and linear in x, and each $h_\ell(x)$ is real and quadratic in x. Thus, solving the matrix problem (2.7) amounts to solving a succession of simpler matrix problems of the form

$$j_\ell(tA)\cdot v = k \quad , \qquad h_\ell(tA)\cdot v = k \ ,$$

for given vectors k. For example, if A is a tridiagonal positive definite Hermitian $N \times N$ matrix, then solving (2.7) reduces to solving a succession of matrix problems for which the matrix involved is either a tridiagonal or five-diagonal positive definite Hermitian $N \times N$ matrix. For problems arising from a two-dimensional (spatial) parabolic partial differential equation, this factorization allows one to use either a direct inversion procedure, or a multi-line successive overrelaxation iterative method.

5. Improved Error Bounds for Spline and L-Spline Interpolation

We now switch to another topic, concerned with approximation by spline and L-spline interpolation. The results of this section are from Swartz and Varga [12]. Fuller details can be found in [12].

We now introduce some standard notation. For $-\infty < a < b < +\infty$, for each integer m and for each extended real number with $1 \leq q \leq \infty$, $W_q^m[a,b]$ denotes the Sobolev space of all real-valued functions $w(x)$ defined on the interval $[a,b]$ such that $w \in C^{m-1}[a,b]$, $D^{m-1}w$

is absolutely continuous with $D^m w \in L_q[a,b]$, where $D \equiv \frac{d}{dx}$. It is well known that $W_q^m[a,b]$ is a Banach space, with its norm defined by

$$(5.1) \qquad \|w\|_{W_q^m[a,b]} = \sum_{j=0}^{m} \|D^j w\|_{L_q[a,b]} .$$

Next, for N a positive integer, let $\Delta: a = x_0 < x_1 < \cdots < x_N = b$ denote a partition Δ of $[a,b]$. The collection of all such partitions Δ of $[a,b]$ is denoted by $\mathcal{P}(a,b)$. We further define $\pi = \max_i (x_{i+1}-x_i)$ and $\underline{\pi} = \min_i (x_{i+1}-x_i)$ for each partition $\Delta \in \mathcal{P}(a,b)$. For any real number B with $B \geq 1$, $\mathcal{P}_B(a,b)$ then denotes the subset of all partitions Δ in $\mathcal{P}(a,b)$ for which $\pi \leq B\underline{\pi}$. In particular, $\mathcal{P}_1(a,b)$ is the collection of all uniform partitions of $[a,b]$, and its elements are denoted by Δ_u.

Since we shall make extensive use of L-splines, we now briefly describe them. Given the differential operator L of order m,

$$(5.2) \qquad Lu(x) \equiv \sum_{j=0}^{m} c_j(x) D^j u(x) , \quad m \geq 1 ,$$

where $c_j \in C^j[a,b]$, $0 \leq j \leq m$, with $c_m(x) \geq \delta > 0$ for all x in $[a,b]$, and given the partition $\Delta: a = x_0 < x_1 < \cdots < x_N = b$, for $N > 1$ let $z = (z_1, \ldots, z_{N-1})$, the <u>incidence vector</u>, be an $(N-1)$-triple of positive integers with $1 \leq z_i \leq m$, $1 \leq i \leq N-1$. Then, $Sp(L,\bar{\Delta},z)$, the L-spline space, is the collection of all real-valued functions on $[a,b]$ such that (cf. Ahlberg, Nilson, and Walsh [1, ch. 6] and Schultz and Varga [11])

$$(5.3)\quad \begin{cases} L^*Lw(x) = 0 \ , \quad x \in [a,b] - \{x_i\}_{i=1}^{N-1} \ , \\ D^k w(x_i-) = D^k w(x_i+) \quad \text{for all} \quad 0 \leq k \leq 2m-1-z_i \ , \\ \qquad\qquad\qquad\qquad\qquad\qquad 0 < i < N \ , \end{cases}$$

where L^* is the formal adjoint of L. From (5.3), we see that $Sp(L,\Delta,z) \subset C^{2m-\sigma-1}[a,b]$ where $\sigma \equiv \max_{1 \leq i \leq N-1} z_i$.

In the special case $L = D^m$, the elements of $Sp(L,\Delta,z)$ are, from (5.3), polynomials of degree $2m-1$ on each subinterval of Δ, and, as such, are called <u>polynomial splines</u>. More specifically, when $L = D^m$ and $z_i = m$, $0 < i < N$, the associated L-spline space is called the <u>Hermite space</u>, and is denoted by $H^{(m)}(\Delta)$. Similarly, when $L = D^m$ and $z_i = 1$, $0 < i < N$, the associated L-spline space is called the <u>spline space</u>, and is denoted by $Sp^{(m)}(\Delta)$.

Finally, if f is any bounded function defined on $[a,b]$, then

$$\omega(f,\delta) \equiv \sup\{|f(x+t) - f(x)| : x,\ x+t \text{ are in } [a,b] \text{ and } |t| \leq \delta\}$$

denotes the usual modulus of continuity of f. In general, K will denote below any generic constant which is independent of the functions considered, and is independent of π.

The following interpolation error bounds are typical (cf. Hedstrom and Varga [4]), and follow from results of Jerome and Varga [5], Schultz and Varga [11], and Perrin [9].

Theorem 5. Given $f \in W_2^{2m}[a,b]$ and given $\Delta \in \mathcal{P}_B(a,b)$, let s be the unique element in $Sp(L,\Delta,z)$ which interpolates f in the sense that

$$(5.4) \qquad D^j(f-s)(x_i) = 0 \, , \quad 0 \leq j \leq z_i - 1 \, , \quad 0 \leq i \leq N \quad (z_o = z_N = m) \, .$$

Then, for $2 \leq q \leq \infty$,

$$(5.5) \qquad \|D^j(f-s)\|_{L_q[a,b]} \leq K\pi^{2m-j-\frac{1}{2}+\frac{1}{q}} \|f\|_{W_2^{2m}[a,b]} \, , \quad 0 \leq j \leq 2m-1 \, .$$

For polynomial splines (i.e., $L = D^m$), $\|f\|_{W_2^2[a,b]}$ can be replaced by $\|D^{2m}f\|_{L_2[a,b]}$ in (5.5).

The above result, based on the second intregal relation (cf. [1, p. 205]), is for rather smooth functions. While the exponent of π in (5.5) is sharp in that it cannot in general be increased for the function spaces considered, our goal is to obtain sharp interpolation errors for less smooth functions f. We next state a result of [12, Lemma 3.2] based on the Peavo Kernel Theorem, which is useful in achieving this goal.

Theorem 6. Given $f \in C^k[a,b]$ with $0 \leq k < 2m$ and given

$\Delta \in \mathcal{P}_B(a,b)$, let g be the unique element in $H^{(2m+1)}(\Delta)$ such that

$$(5.6)\qquad \begin{cases} D^j(f-g)(x_i) = 0 , & 0 \le j \le k , \quad 0 \le i \le N , \\ D^j g(x_i) = 0 , & k < j \le 2m , \quad 0 \le i \le N . \end{cases}$$

Then,

$$(5.7)\qquad K\pi^{k-j}\omega(D^k f,\pi) \ge \begin{cases} \|D^j(f-g)\|_{L_\infty[a,b]} , & 0 \le j \le k , \\ \|D^j g\|_{L_\infty[a,b]} , & k < j \le 2m . \end{cases}$$

With Theorem 6, we now prove an analogue of Theorem 5 for less smooth functions.

<u>Theorem 7</u>. Given $f \in C^k[a,b]$ with $0 \le k < 2m$ and given $\Delta \in \mathcal{P}_B(a,b)$, let s be the unique element in $Sp(L,\Delta,z)$ such that for $z_0 = z_N = m$,

$$(5.8)\qquad \begin{cases} D^j(f-s)(x_i) = 0 , & 0 \le j \le \min(k,z_i-1), \quad 0 \le i \le N, \\ D^j s(x_i) = 0 & , \text{ if } \min(k,z_i-1) < j \le z_i-1 , \\ & \qquad 0 \le i \le N . \end{cases}$$

Then, for $2 \le q \le \infty$,

$$(5.9)\qquad K\left\{\pi^{k-j-\frac{1}{2}+\frac{1}{q}}\,\omega(D^k f,\pi) + \pi^{2m-j-\frac{1}{2}+\frac{1}{q}}\,\|f\|_{W_2^k[a,b]}\right\}$$

$$\ge \begin{cases} \|D^j(f-s)\|_{L_q[a,b]} , & 0 \le j \le k , \\ \|D^j s\|_{L_q[a,b]} , & \text{if } k < j \le 2m-1 . \end{cases}$$

For polynomial splines $(L = D^m)$, the term involving $\|f\|_{W_2^k[a,b]}$ can be deleted in (5.9).

Proof. Given $f \in C^k[a,b]$, let g be its interpolant in $H^{(2m+1)}(\Delta)$, in the sense of (5.6) of Theorem 6. For any $2 \leq q \leq \infty$, the triange inequality gives us that

$$(5.10) \qquad \|D^j(f-s)\|_{L_q[a,b]} \leq \|D^j(f-g)\|_{L_q[a,b]} + \|D^j(g-s)\|_{L_q[a,b]}, \quad 0 \leq j \leq k,$$

where s is the interpolant of f is $Sp(L,\Delta,z)$ in the sense of (5.8). Note from (5.8) that s is also the interpolant of g in $Sp(L,\Delta,z)$ in the sense of (5.4). Hence, applying Theorem 5 yields

$$(5.11) \quad \|D^j(g-s)\|_{L_q[a,b]} \leq K\pi^{2m-j-\frac{1}{2}+\frac{1}{q}} \|g\|_{W_2^{2m}[a,b]}, \quad 0 \leq j \leq 2m-1.$$

We now bound $\|g\|_{W_2^{2m}[a,b]}$. For any ℓ with $k < \ell \leq 2m$, (5.7) of Theorem 6 gives

$$(5.12) \qquad \|D^\ell g\|_{L_2[a,b]} \leq K\pi^{k-\ell}\,\omega(D^k f,\pi), \quad k < \ell \leq 2m .$$

For any ℓ with $0 \leq \ell \leq k$, we evidently have

$$\|D^\ell g\|_{L_2[a,b]} \leq \|D^\ell(f-g)\|_{L_2[a,b]} + \|D^\ell f\|_{L_2[a,b]}, \quad 0 \leq \ell \leq k,$$

and, using the first inequality of (5.7) of Theorem 6, this can be bounded above by

$$(5.13)\quad \|D^{\ell}g\|_{L_2[a,b]} \leq K\pi^{k-\ell}\,\omega(D^k f,\pi) + \|D^{\ell}f\|_{L_2[a,b]}\ , \qquad 0 \leq \ell \leq k\ .$$

Summing the inequalities of (5.12) and (5.13) and using the norm definition of (5.1) gives

$$\|g\|_{W_2^{2m}[a,b]2} \leq K\left\{\pi^{k-2m}\,\omega(D^k f,\pi) + \|f\|_{W_2^k[a,b]}\right\}\ .$$

This bound, when substituted in (5.11), gives

$$\|D^j(g-s)\|_{L_q[a,b]} \leq K\left\{\pi^{k-j-\frac{1}{2}+\frac{1}{q}}\,\omega(D^k f,\pi) + \pi^{2m-j-\frac{1}{2}+\frac{1}{q}}\,\|f\|_{W_2^k[a,b]}\right\}$$

for $0 \leq j \leq 2m-1$, thus suitably bounding the last term of (5.10). If polynomial splines are used, the term involving $\|f\|_{W_2^k[a,b]}$ can be deleted (cf. Theorem 5). Finally, the first term of the right side of (5.10) can be bounded above from (5.7) of Theorem 6, and the combined upper bounds, when inserted in (5.10), give the desired result of the first inequality of (5.9) for $0 \leq j \leq k$. If $k < j \leq 2m-1$, the same technique can be used to bound the terms on the right-hand side of

$$\|D^j s\|_{L_q[a,b]} \leq \|D^j(g-s)\|_{L_q[a,b]} + \|D^j g\|_{L_q[a,b]}\ , \qquad k < j \leq 2m-1,$$

which then establishes the second inequality of (5.9) for $k < j \leq 2m-1$. Q.E.D.

As an easy extension of Theorem 7, we include (cf. [12])

Corollary 1. With the hypotheses of Theorem 7, if $f \in W_r^{k+1}[a,b]$ with $1 \leq r \leq \infty$ and $0 \leq k < 2m$, then for $\max(r,2) \leq q \leq \infty$,

$$(5.14) \qquad K\pi^{k+1-j+\frac{1}{q}+\min(-\frac{1}{r},-\frac{1}{2})} \|f\|_{W_r^{k+1}[a,b]} \geq \begin{cases} \|D^j(f-s)\|_{L_q[a,b]} , & 0 \leq j \leq k , \\ \|D^j s\|_{L_q[a,b]} , & \text{if } k < j \leq 2m-1 . \end{cases}$$

One difficulty with the result of Theorem 7 is that one needs to know the explicit continuity class of f to define its interpolant s in $Sp(L,\Delta,z)$ in the sense of (5.8). Often, this continuity class is difficult to determine from raw data in a routine setting in, say, a computer center. However, this can be avoided through the use of Lagrange interpolation. We now state (cf. [12])

Theorem 8. Given $f \in C^k[a,b]$ with $0 \leq k < 2m$ and given $\Delta \in \mathcal{P}_B(a,b)$ with at least $2m$ knots, let s be the unique element in $Sp(L,\Delta,z)$ such that for $z_0 = z_N = m$,

$$(5.15) \qquad D^j s(x_i) = D^j(\mathcal{L}_{2m-1,i}f)(x_i) , \quad 0 \leq j \leq z_i-1, \quad 0 \leq i \leq N ,$$

where $\mathcal{L}_{2m-1,i}f$ is any Lagrange polynomial (of degree $2m-1$) interpolation of f in $2m$ consecutive knots $x_j, x_{j+1}, \ldots, x_{j+2m-1}$ with $x_i \in [x_j, x_{j+2m-1}]$. Then, for

$2 \leq q \leq \infty$, the bounds of (5.9) are valid, where again, for polynomial splines, the term in (5.9) involving $\|f\|_{W_2^k[a,b]}$ can be deleted.

We remark that the inequalities of (5.14) of Corollary 1 also apply to the interpolant s in $Sp(L,\Delta,z)$, defined by (5.15).

Although the result of Theorem 8 lifts the objection raised concerning the application of Theorem 7, we note that, for polynomial splines and $q = \infty$, the inequalities of (5.9) become

$$K\pi^{k-j-\frac{1}{2}}\omega(D^k f,\pi) \geq \begin{cases} \|D^j(f-s)\|_{L_\infty[a,b]}, & 0 \leq j \leq k, \\ \|D^j s\|_{L_\infty[a,b]}, & \text{if } k < j \leq 2m-1, \end{cases}$$

and one naturally expects that the exponent of π in the above expression is too small by a factor $\frac{1}{2}$. To improve the inequalities above, we next state another result of [12].

<u>Theorem 9</u>. Given $f \in C^k[a,b]$ with $0 \leq k < 2m$ and given $\Delta_u \in \mathcal{P}_1(a,b)$ with at least $2m$ knots, let s be the unique element in $Sp^{(m)}(\Delta_u)$ (i.e., $z_1 = z_2 = \cdots = z_{N-1} = 1$) such that

$$(5.16) \qquad \begin{cases} (f-s)(x_i) = 0, & 0 \leq i \leq N, \\ D^j(f-s)(a) = D^j(f-s)(b) = 0, & 0 \leq j \leq \min(k,m-1), \\ D^j s(a) = D^j s(b) = 0, & \text{if } \min(k,m-1) < j \leq m-1, \end{cases}$$

or

$$(5.17)\qquad \begin{cases} (f-s)(x_i) = 0 , & 0 \leq i \leq N , \\ D^j(f-s)(a) = D^j(\mathscr{L}_{2m-1,0}f)(a) , & 0 \leq j \leq m-1 , \end{cases}$$

where $\mathscr{L}_{2m-1,0}f$ is, as in Theorem 8, the Lagrange polynomial interpolation of f in the knots $x_0, x_1, \ldots, x_{2m-1}$, with a similar definition at $x = b$. Then,

$$(5.18)\qquad K\pi^{k-j}\,\omega(D^k f,\pi) \geq \begin{cases} \|D^j(f-s)\|_{L_\infty[a,b]} , & 0 \leq j \leq k , \\ \|D^j s\|_{L_\infty[a,b]} , & \text{if } k < j \leq 2m-1 . \end{cases}$$

Corollary 2. With the hypotheses of Theorem 9, if $f \in W_r^{k+1}[a,b]$ with $0 \leq k < 2m$ and $1 \leq r \leq \infty$, then for $\max(r,2) \leq q \leq \infty$,

$$(5.19)\qquad K\pi^{k+1-j-\frac{1}{r}+\frac{1}{q}} \|D^{k+1}f\|_{L_r[a,b]} \geq \begin{cases} \|D^j(f-s)\|_{L_q[a,b]} , & 0 \leq j \leq k , \\ \|D^j s\|_{L_q[a,b]} , & \text{if } k < j \leq 2m-1 . \end{cases}$$

To give an explicit example of Theorem 9, consider the particular case $m = 2$ of cubic splines. The cubic spline $s \in Sp^{(2)}(\Delta_u)$ of (5.17) is defined then by

$$(f-s)(x_i) = 0 , \quad 0 \leq i \leq N ,$$

$$Ds(a) = \frac{1}{6h}\left\{-11f(a) + 18f(a+h) - 9f(a+2h) + 2f(a+3h)\right\} ,$$

$$Ds(b) = \frac{1}{6h}\left\{11f(b) - 18f(b-h) + 9f(b-2h) - 2f(b-3h)\right\} ,$$

where $h = \pi$ measures the uniform mesh of Δ_u. In this special case, (5.19) becomes, for $f \in C^k[a,b]$ with $0 \leq k < 4$,

$$Kh^{k-j}\,\omega(D^k f,h) \geq \begin{cases} \|D^j(f-s)\|_{L_\infty[a,b]}, & 0 \leq j \leq k, \\ \|D^j s\|_{L_\infty[a,b]}, & \text{if } k < j \leq 3. \end{cases}$$

More complete results can be found in [12] for Hermite L-splines. In addition, certain stability theorems are established in [12] concerning the use of Lagrange interpolation of f to define interpolants s in $Sp(L,\Delta,z)$.

6. Application to Two-Point Boundary Value Problems

The interpolation error bounds of Section 5 can be applied to the numerical solution of two-point nonlinear boundary value problems in the following way. Consider, as a very special case, the Galerkin approximation of the solution of

$$\text{(6.1)} \qquad \begin{cases} (-1)^{m+1} D^{2m} u(x) = f(x,u(x)), & a < x < b, \\ D^j u(a) = D^j u(b) = 0, & 0 \leq j \leq m-1, \end{cases}$$

where it is assumed that $f(x,u)$ is a real-valued function defined on $[a,b] \times R$ such that $f(x,u)$ and $f_u(x,u)$ are in $C^\circ([a,b] \times R)$, and there exists a constant γ such that

$$\text{(6.2)} \qquad f_u(x,u) \geq \gamma > -\Lambda \quad \text{for all } x \in [a,b], \text{ and all real } u,$$

where $\Lambda \equiv \inf\left\{\int_a^b (D^m w(x))^2 dx / \int_a^b (w(x))^2 dx : w \in \overset{\circ}{W}{}_2^m[a,b]\right\}$,

and where $\overset{\circ}{W}{}_2^m[a,b]$ denotes the subspace of functions of $W_2^m[a,b]$ which satisfy the boundary conditions of (6.1). Given a finite dimensional subspace S_M of $\overset{\circ}{W}{}_2^m[a,b]$ with $\{w_i(x)\}_{i=1}^M$ a basis for S_M , the Galerkin approximation $\hat{w}(x) \in S_M$ of the solution of (6.1) is characterized (cf. [10]) by

$$(6.3) \quad \int_a^b \left\{D^m\hat{w}(t)\cdot D^m w_i(t) + f(t,\hat{w}(t))\cdot w_i(t)\right\}dt = 0 , \quad 1 \leq i \leq M .$$

Next, let $\overset{\circ}{Sp}{}^{(m)}(\Delta_u)$ denote the subspace of functions in $Sp^{(m)}(\Delta_u)$ which satisfy the boundary conditions of (6.1). Then, based on results of Perrin, Price, and Varga [10, Theorem 3] and Theorem 9 of Section 5, we have

Theorem 10. If $u(x)$, the generalized solution of (6.1) is in $C^{2m}[a,b]$, and $\hat{w}(x)$ is its unique Galerkin approximation in $\overset{\circ}{Sp}{}^{(m)}(\Delta_u)$, then

$$(6.4) \quad \|D^j(u-w)\|_{L_\infty[a,b]} \leq K\pi^{2m-j} \|D^{2m}u\|_{L_\infty[a,b]} , \quad 0 \leq j \leq m-1 .$$

REFERENCES

1. Ahlberg, J. H., Nilson, E. N., and Walsh, J. L., The Theory of Splines and their Applications, Academic Press, New York, 1967.

2. Boas, R. P. Jr., Entire Functions, Academic Press, New York, 1954.
3. Cody, W. J., Meinardus, G., and Varga, R. S., "Chebyshev Rational Approximations to e^{-x} in $[0,+\infty)$ and Applications to Heat-Conduction Problems," J. Approx. Theory, 2, 50-65 (1969).
4. Hedstrom, Gerald W., and Varga, Richard S., "Application of Besov Spaces to Spline Approximation," to appear.
5. Jerome, J. W., and Varga, R. S., "Generalizations of Spline Functions and Applications to Nonlinear Boundary Value and Eigenvalue Problems," Theory and Applications of Spline Functions (edited by T. N. E. Greville), Academic Press, New York, 1969, 103-155.
6. Meinardus, Günter, Approximation of Functions: Theory and Numerical Methods, Springer-Verlag, New York, 1967. Translated by L. L. Schumaker.
7. Meinardus, G., Taylor, G., Reddy, A. R., and Varga, R. S., "Converse Theorems and Extensions in Chebyshev Rational Approximation in $[0,+\infty)$," to appear.
8. Meinardus, Günter, and Varga, Richard S., "Chebyshev Rational Approximation to Certain Entire Functions in $[0,+\infty)$," to appear.
9. Perrin, F. M., "An Application of Monotone Operators to Differential and Partial Differential Equations on Infinite Domains," Doctoral Thesis, Case Institute of Technology, 1967.
10. Perrin, F. M., Price, H. S., and Varga, R. S., "On Higher-Order Numerical Methods for Nonlinear Two-Point Boundary Value Problems," Numer. Math. 13 (1969), 180-198.
11. Schultz, M. H., and Varga, R. S., "L-Splines," Numer. Math. 10 (1967), 345-369.
12. Swartz, Blair, and Varga, Richard S., "Error Bounds for Spline and L-Spline Interpolation," to appear.
13. Valiron, G., Lectures on the General Theory of Integral Functions, Chelsea, New York, 1949.
14. Varga, Richard S., Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.