JOURNAL OF APPROXIMATION THEORY **13**, 128–135 (1975)

Minimum Moduli of Differential Operators from the Viewpoint of Approximation Theory

RONALD A. DEVORE*

Department of Mathematics, Oakland University, Rochester, Michigan 48063

Communicated by P. L. Butzer

DEDICATED TO PROFESSOR G. G. LORENTZ ON THE OCCASION OF HIS SIXTY-FIFTH BIRTHDAY

1. INTRODUCTION

One of the prettiest results in approximation theory is an old theorem of S. Bernstein which states that if $f^{(n)}$ is absolutely continuous on $[-1, 1]$ and $f^{(n+1)}$ is in $L^\infty[-1, 1]$, then

$$\operatorname{dist}_\infty(f, \mathbf{P}_n) \leqslant \frac{2^{-n}}{(n+1)!} \| f^{(n+1)} \|_\infty , \tag{1}$$

where the distance is measured in the $L^\infty[-1, 1]$ norm, $\| \cdot \|_\infty$, and $\mathbf{P}_n$ is the space of algebraic polynomials of degree $\leqslant n$. The inequality (1) is sharp, since when $f(x) = 2^{-n}/(n+1)! \cos(n+1) \arccos x$, it becomes an equality. In approximation theoretic terms, (1) provides an estimate for the error in approximating f in $\| \cdot \|_\infty$ by elements of $\mathbf{P}_n$. Of course, it is of interest to obtain results for other norms (e.g., L^p) and other spaces. This paper makes a modest contribution in this direction.

In order to provide the proper setting for our generalizations, we introduce the idea of the minimum modulus of a differential operator. Suppose T is an ordinary linear differential operator of order n with domain $\mathscr{D}(T) =: \{f \in L^p[-1, 1]: f^{(n-1)} \text{ a.c.}, Tf \in L^q[-1, 1]\}$. Denote by $\mathscr{N}(T)$, the null space of T. When considered as a mapping on $\mathscr{D}(T)/\mathscr{N}(T)$ into $L^q[-1, 1]$, T has an inverse. The reciprocal of the norm of this inverse operator is called the minimum modulus of T, $\gamma(T, p, q)$. It is easy to verify the formula

$$(\gamma(T, p, q))^{-1} = \sup_{\|Tf\|_q \leqslant 1} \operatorname{dist}_p(f, \mathscr{N}(T)). \tag{2}$$

* The author gratefully acknowledges NSF support in Grant GP19620.

If there is a function f in $\mathscr{D}(T)$ for which the supremum on the right-hand side of (2) is attained, we say f is extremal.

We can now restate (1) in terms of the minimum modulus. We use $n + 1$ instead of n and consider the operator $T = D^{n+1}$. The null space of T is $\mathbf{P}_n$. Hence, the inequality (1) and the fact that it is sharp is exactly the same as saying

$$\gamma(D^{n+1}, \infty, \infty) = 2^n(n + 1)!\,.$$

Also, the function $2^{-n}/(n + 1)! \cos(n + 1) \arccos x$ is extremal.

In this sense, we see that (1) is really just the determination of $\gamma(D^{n+1}, \infty, \infty)$. It is in this spirit that we seek generalizations of (1). Thus, we will replace D^{n+1} by more general operators T, and replace $\mathbf{P}_n$ by $\mathscr{N}(T)$. We would also like to replace L^∞ by other L^p spaces. The problem then is to determine $\gamma(T, p, q)$. When this is accomplished, we have the inequality

$$\operatorname{dist}_p(f, \mathscr{N}(T)) \leqslant (\gamma(T, p, q))^{-1} \| Tf \|_q\,, \tag{3}$$

as our generalization to (1), and of course (3) is sharp.

Our techniques will be applicable to operators T of the form

$$T = (D + \lambda_n(x)) \cdots (D + \lambda_1(x)), \qquad \lambda_k \in C^{(n-k)}[-1, 1]. \tag{4}$$

The functions λ_k are assumed to be real valued. These requirements on T guarantee, among other things, that $\mathscr{N}(T)$ is a Chebyshev space of dimension n.

In Section 2, we will determine $\gamma(T, p, \infty)$, for each $1 \leqslant p \leqslant \infty$. The case $\gamma(T, \infty, \infty)$ is a result of M. Zedek [7]. The value of $\gamma(T, \infty, \infty)$ was also obtained by T. Rivlin [6] with a different point of view. Our approach differs from Rivlin's and Zedek's. It is more in line with the traditional proofs of Bernstein's inequality. In Section 3, we will determine $\gamma(T, 1, 1)$. Here, our approach is approximation theoretic, relying heavily on duality and characterizations of best approximations in L^1.

The reader will find that conspicuously absent is the determination of $\gamma(T, 2, 2)$, which on the surface would appear to be the most manageable because of all the structure in L^2. This case can be handled in a theoretical sense using a calculus of variations approach as is done in the paper of S. Goldberg and A. Meir [3]. However, the determination of the numerical value of $\gamma(T, 2, 2)$, even for $T = D^n$, appears to be a formidable problem.

2. The Determination of $\gamma(T, p, \infty)$

When T is a differential operator of the form (4), then $\mathscr{N}(T)$ is a Chebyshev space of dimension n (Zedek [7]). This means that we can interpolate any

n values by functions in $\mathcal{N}(T)$. We will use a remainder formula for this interpolation which is the analogue of Cauchy's formula for Lagrange interpolation.

LEMMA 1. *If $f \in C^{(n)}[-1, 1]$ has $(n+1)$ distinct zeros in $[-1, 1]$, then there is a point ξ in $(-1, 1)$ for which $Tf(\xi) = 0$.*

Proof. This is a result of Zedek [7].

LEMMA 2. *Suppose ψ is a solution to the equation $T\psi = 1$ on $[-1, 1]$ which has exactly n distinct zeros $x_1, x_2, \ldots, x_n$ in $[-1, 1]$. If $f \in C^{(n)}[-1, 1]$ and $P \in \mathcal{N}(T)$ interpolates f at each point $x_1, x_2, \ldots, x_n$, i.e., $P(x_i) = f(x_i)$, $i = 1, 2, \ldots, n$, then for each $x \in [-1, 1]$ there is a $\xi_x \in (-1, 1)$ such that*

$$f(x) - P(x) = Tf(\xi_x)\,\psi(x). \tag{5}$$

Proof. The proof is an exact mimic of the proof of Cauchy's formula for $T = D^n$. The formula (5) is clear when x is one of the points $x_1, x_2, \ldots, x_n$. When $x \neq x_i$, $i = 1, 2, \ldots, n$, let $\alpha = (\psi(x))^{-1} \cdot (f(x) - P(x))$. The function $f(t) - P(t) - \alpha\psi(t)$ vanishes at the $n+1$ distinct points $x, x_1, x_2, \ldots, x_n$, and hence by Lemma 1 there is a point ξ_x for which $T(f - P - \alpha\psi)(\xi_x) = 0$. Since, $TP = 0$ and $T\psi = 1$ on $[-1, 1]$, the last equation can be rewritten as $\alpha = Tf(\xi_x)$ which gives (5) by the very definition of α.

Now, let ψ be any function in $C^{(n)}[-1, 1]$ for which $T\psi = 1$ on $[-1, 1]$. We want to approximate ψ by elements of $\mathcal{N}(T)$ in the $L^p[-1, 1]$ norm. The existence and uniqueness of such approximants are classical results [4]. Denote by P_p^*, the best $L^p[-1, 1]$ approximation to ψ from $\mathcal{N}(T)$, so that

$$\| \psi - P_p^* \|_p = \inf_{P \in \mathcal{N}(T)} \| \psi - P \|_p .$$

LEMMA 3. *For each $1 \leq p \leq \infty$, the function $\psi_p = \psi - P_p^*$ has exactly n distinct zeros in $[-1, 1]$ and changes sign at each of these zeros.*

Proof. We first show that ψ_p has at least n changes of sign in $[-1, 1]$. For $p = 1, \infty$, this follows from the classical alternation theorems. For $1 < p < \infty$, the proof is simple enough. By the duality theorem for L^p approximation [4, p. 84], the function $h_p = |\psi_p|^{p-1} \operatorname{sgn} \psi_p$ is orthogonal to $\mathcal{N}(T)$, i.e.,

$$\int_{-1}^{1} P(x)\, h_p(x)\, dx = 0, \qquad P \in \mathcal{N}(T). \tag{6}$$

Notice that h_p changes sign precisely at the points where ψ_p changes sign. If ψ_p and hence h_p has less than n changes of sign in $(-1, 1)$, then we can construct a function $P \in \mathcal{N}(T)$ which changes sign precisely at these points (this is a well-known property of Chebyshev systems [5, p. 30]). This makes it clear that

$$\int_{-1}^{1} P(x)\, h_p(x)\, dx \neq 0,$$

which is the desired contradiction. Thus, ψ_p has at least n changes of sign in $(-1, 1)$.

That ψ_p can have no more than n zeros in $[-1, 1]$ follows from Lemma 1. For if not there would be a point $\xi \in (-1, 1)$, with $T\psi_p(\xi) = 0$ which contradicts the fact that $T\psi_p = 1$ on $[-1, 1]$.

We can now easily prove the main result of this section.

THEOREM 1. *If T is of the form* (4), *then for* $1 \leqslant p \leqslant \infty$, $\gamma(T, p, \infty) = \| \psi_p \|_p^{-1}$ *and* ψ_p *is an extremal function.*

Proof. Let $x_1, x_2, \ldots, x_n$ be the n distinct zeros of ψ_p in $[-1, 1]$. Suppose first that $f \in C^{(n)}[-1, 1]$ and $P_f \in \mathcal{N}(T)$ is the function which interpolates f at $x_1, x_2, \ldots, x_n$. From Lemma 2, we see that if $x \in [-1, 1]$, there is a point $\xi_x \in (-1, 1)$ such that

$$f(x) - P_f(x) = Tf(\xi_x)\, \psi_p(x).$$

So,

$$|f(x) - P_f(x)| \leqslant \| Tf \|_\infty \, | \psi_p(x) |, \qquad -1 \leqslant x \leqslant 1.$$

Taking the norm in L^p, we find

$$\| f - P_f \|_p \leqslant \| Tf \|_\infty \| \psi_p \|_p ,$$

or

$$\operatorname{dist}_p(f, \mathcal{N}(T)) \leqslant \| \psi_p \|_p \| Tf \|_\infty . \tag{7}$$

This result also holds for all f with $f^{(n-1)}$ absolutely continuous and $Tf \in L^\infty[-1, 1]$, because of the denseness of $C^{(n)}[-1, 1]$ in this space. The estimate (7) shows that $\gamma(T, p, \infty) \geqslant \| \psi_p \|^{-1}$. The opposite inequality is immediate when we take $f = \psi_p$.

When $T = D^n$, the functions ψ_p and the values of $\gamma(D^n, p, \infty)$ are easily obtained for $p = 1, 2, \infty$. For $p = 1$, ψ_1 is the normalized Chebyshev polynomial of the second kind and $\gamma(D^n, 1, \infty) = 2^{n-1} n!$. For $p = 2$, ψ_2 is the normalized Legendre polynomial of degree n and $\gamma(D^n, 2, \infty) = (n + 1/2)^{1/2} (2n)!/2^n n!$. For $p = \infty$, ψ_∞ is the Chebyshev polynomial of degree n and $\gamma(D^n, \infty, \infty) = 2^{n-1} n!$, which is again Bernstein's result (1).

3. The Determination of $\gamma(T, 1, 1)$

In this case the situation is a little more subtle, mainly because there is no extremal function in $\mathscr{D}(T)$. We will use a well-known duality theorem for L^1 approximation which we now state.

Lemma 4. *Let $f \in L^1[-1, 1]$ and $\mathscr{M}$ be a finite dimensional subspace of $L^1[-1, 1]$. Then,*

$$\operatorname{dist}_1(f, \mathscr{M}) = \sup_{h \in \mathscr{M}^\perp} \int_{-1}^{1} f(x)\, h(x)\, dx,$$

where $\mathscr{M}^\perp = \{h \in L^\infty : \| h \|_\infty = 1$ and $\int_{-1}^{1} P(x)\, h(x)\, dx = 0,\ P \in \mathscr{M}\}$.

Recall, the function ψ_1 introduced in the last section. Since the span of $\mathscr{N}(T) \cup \{\psi_1\}$ is a Chebyshev space (use Lemma 1), we also have available Markov's theorem for L^1 approximation [4, p. 67].

Lemma 5. *Let $f \in C[-1, 1]$ and $P_f \in \mathscr{N}(T)$ be that function which interpolates f at $x_1, x_2, \ldots, x_n$ the zeros of ψ_1. If $f - P_f$ changes sign precisely at $x_1, x_2, \ldots, x_n$, then P_f is the best $L^1[-1, 1]$ approximation to f from $\mathscr{N}(T)$ and*

$$\operatorname{dist}_1(f, \mathscr{N}(T)) = \left| \int_{-1}^{1} f(x) \operatorname{sgn} \psi_1(x)\, dx \right|.$$

Lemma 6. *If $f \in C^{(n)}[-1, 1]$ with $Tf > 0$ ($Tf < 0$) on $[-1, 1]$, then P_f is the best $L^1[-1, 1]$ approximation to f from $\mathscr{N}(T)$ and*

$$\operatorname{dist}_1(f, \mathscr{N}(T)) = \left| \int_{-1}^{1} f(x) \operatorname{sgn} \psi_1(x)\, dx \right|.$$

Proof. From Lemma 2, we see that

$$f(x) - P_f(x) = Tf(\xi_x)\, \psi_1(x).$$

Since $Tf > 0$, $f - P_f$ will change sign precisely when ψ_1 does and thus the result follows from Lemma 5.

We can now determine $\gamma(T, 1, 1)$. Let T^* be the adjoint operator to T (see [2, p. 1285]). Denote by Ψ_1, the solution to the differential equation $T^*\Psi = \operatorname{sgn} \psi_1$, with initial conditions $\Psi_1^{(k)}(-1) = 0$, $k = 1, 2, \ldots, n$. When $f \in C^{(n)}[-1, 1]$ and $P \in \mathscr{N}(T)$ with $P^{(k)}(1) = f^{(k)}(1)$, $k = 1, 2, \ldots, n$, then

$$\int_{-1}^{1} f(x) \operatorname{sgn} \psi_1(x)\, dx = \int_{-1}^{1} (f(x) - P(x)) \operatorname{sgn} \psi_1(x)\, dx$$
$$= \int_{-1}^{1} (f(x) - P(x))\, T^*\Psi_1(x)\, dx$$
$$= \int_{-1}^{1} T(f - P)(x)\, \Psi_1(x)\, dx$$
$$= \int_{-1}^{1} Tf(x)\, \Psi_1(x)\, dx. \tag{8}$$

Similarly, if h is any function in $\mathscr{N}(T)^\perp$ and H satisfies $T^*H = h$, with $H^{(k)}(-1) = 0$, $k = 1, 2,..., n$, then

$$\int_{-1}^{1} f(x)\, h(x)\, dx = \int_{-1}^{1} Tf(x)\, H(x)\, dx \tag{9}$$

THEOREM 2. *If T is of the form* (4), *then* $\gamma(T, 1, 1) = \| \Psi_1 \|_\infty^{-1}$.

Proof. First, let $f_0 \in C^{(n)}[-1, 1]$ with $\| Tf_0 \|_1 = 1$ and $h \in \mathscr{N}(T)^\perp$. Then, by (9)

$$\int_{-1}^{1} f_0(x)\, h(x)\, dx = \int_{-1}^{1} Tf_0(x)\, H(x)\, dx \leqslant \sup_{\|Tf\|_1=1} \int_{-1}^{1} Tf(x)\, H(x)\, dx. \tag{10}$$

It is to be understood in (10) and the sequel that the suprema are taken only over functions in $C^{(n)}[-1, 1]$, unless explicitly stated otherwise. Now, the supremum in (10) is the $L^\infty[-1, 1]$ norm of H. If $\| H \|_\infty = H(x_0)$ for some $x_0 \in [-1, 1]$, then the supremum can be attained by considering only functions with $Tf > 0$. Hence, in this case,

$$\int_{-1}^{1} f_0(x)\, h(x)\, dx \leqslant \sup_{\substack{Tf>0 \\ \|Tf\|_1=1}} \int_{-1}^{1} Tf(x)\, H(x)\, dx$$
$$= \sup_{\substack{Tf>0 \\ \|Tf\|_1=1}} \int_{-1}^{1} f(x)\, h(x)\, dx \leqslant \sup_{\substack{Tf>0 \\ \|Tf\|_1=1}} \left| \int_{-1}^{1} f(x) \operatorname{sgn} \psi_1(x)\, dx \right|$$
$$= \sup_{\substack{Tf>0 \\ \|Tf\|_1=1}} \left| \int_{-1}^{1} Tf(x)\, \Psi_1(x)\, dx \right| \leqslant \| \Psi_1 \|_\infty .$$

Here, for the second inequality, we used Lemmas 6 and 4.

Similarly, when $\| H \|_\infty = -H(x_0)$, we need only consider f's with $Tf < 0$. Arguing as we have above, we find that

$$\int_{-1}^{1} f_0(x)\, h(x)\, dx \leqslant \| \Psi_1 \|_\infty ,$$

for all $h \in \mathcal{N}(T)$. Hence, from Lemma 4, it follows that

$$\operatorname{dist}_1(f_0, \mathcal{N}(T)) \leqslant \| \Psi_1 \|_\infty, \tag{11}$$

whenever $f_0 \in C^{(n)}[-1, 1]$, with $\| Tf_0 \|_1 = 1$. The restriction $f_0 \in C^{(n)}$ is removed by a denseness argument. Thus, (11) shows that

$$\gamma(T, 1, 1) \geqslant \| \Psi_1 \|_\infty^{-1}. \tag{12}$$

To see the reverse inequality, consider any $f \in C^{(n)}[-1, 1]$, with $Tf > 0$ ($Tf < 0$) and $\| Tf \|_1 = 1$. Then, from Lemma 6

$$\operatorname{dist}_1(f, \mathcal{N}(T)) = \left| \int_{-1}^{1} f(x) \operatorname{sgn} \psi_1(x)\, dx \right| = \left| \int_{-1}^{1} Tf(x)\, \Psi_1(x)\, dx \right|.$$

Taking a supremum over all such f we see that the right-hand side becomes the $L^\infty[-1, 1]$ norm of Ψ_1, so that

$$\sup_{\|Tf\|_1 = 1} \operatorname{dist}_1(f, \mathcal{N}(T)) \geqslant \| \Psi_1 \|_\infty .$$

In other words,

$$\gamma(T, 1, 1) \leqslant \| \Psi_1 \|_\infty^{-1}.$$

This is the reverse inequality to (12) and proves the theorem.

When we take $T = D^n$, the function ψ_1 is the Chebyshev polynomial of the second kind of degree n. Hence $\operatorname{sgn} \psi_1 = \operatorname{sgn} \sin(n + 1) \arccos x$. This means that $\operatorname{sgn} \psi_1$ changes sign at the points $\cos(k\pi/(n + 1))$, $k = 1, 2, \ldots, n$. The points $\cos(k\pi/(n + 1))$ are spaced so that the distance between consecutive points increases as we move from -1 to 0 and decreases as we move from 0 to 1. Because of this, an induction argument shows that $\| \Psi_1 \|_\infty$ is $| \Psi_1(0) |$ when n is odd and $| \Psi_1(\cos(n + 2)\pi/(2n + 2)) |$ when n is even.

Rather than try to determine $\| \Psi_1 \|_\infty$ directly, it is easier to return to the ideas used in the proof of Theorem 2. Consider the case when $\| \Psi_1 \|_\infty = \Psi_1(0)$. Then,

$$\Psi_1(0) = \int_{-1}^{1} \Psi_1(x)\, d\mu(x),$$

where $d\mu$ is the Dirac measure with unit mass at 0. The measure $d\mu$ is not the nth derivative of a function from $\mathscr{D}(D^n)$ which is why we dont have an extremal function. However,

$$\begin{aligned} \Psi_1(0) &= \int_{-1}^{1} \Psi_1(x)\, d\mu(x) \\ &= \frac{1}{(n-1)!} \int_{-1}^{1} x_+^{n-1} \operatorname{sgn} \psi_1(x)\, dx = \frac{1}{(n-1)!} \operatorname{dist}_1(x_+^{n-1}, \mathbf{P}_{n-1}), \end{aligned}$$

where x_+^{n-1} is defined to be 0 if $x < 0$ and x^{n-1} if $x > 0$.

Even though there is no extremal function in the strict sense (x_+^{n-1} is not in $\mathscr{D}(D^n)$), the function $x_+^{n-1}/(n-1)!$ still serves the purpose of determining $\gamma(D^n, 1, 1)$ when n is odd. Similarly,

$$(\gamma(D^n, 1, 1))^{-1} = \frac{1}{(n-1)!} \operatorname{dist}_1((x\text{-}\cos(n+2)\,\pi/(2n+2))_+^{n-1}, \mathbf{P}_{n-1}),$$

when n is even. The problem of determining $\operatorname{dist}_1(x_+^{n-1}, \mathbf{P}_{n-1})$ is solved explicitly in [1] by means of a finite but complicated sum which we do not reproduce here. When n is even, the results of [1] do not determine $\gamma(D^n, 1, 1)$ explicitly but do provide asymptotic estimates.

References

1. R. DeVore, On L_p-approximation of functions whose mth derivative is of bounded variation, *Studia Sci. Math. Hung.* **3** (1968), 443–450.
2. N. Dunford and J. Schwartz, "Linear Operators," Part II, Interscience, New York, 1963.
3. S. Goldberg and A. Meir, Minimum moduli of ordinary differential operators, preprint.
4. M. Golomb, "Lectures on the Theory of Approximation," Argonne National Laboratory, 1962.
5. S. Karlin and W. Studden, "Tchebycheff Systems: With Applications in Analysis and Statistics," Interscience, New York, 1966.
6. T. Rivlin, A duality theorem and upper bounds for approximation, *in* "Abstract Spaces and Approximation" (P. L. Butzer and B. Sz.-Nagy, Eds.), pp. 274–280, Proceedings of the conference held in Oberwolfach, 1968, Birkhäuser Verlag, Basel, 1969.
7. M. Zedek, On approximation by solutions of ordinary linear differential equations, *SIAM J. Numer. Anal.* **3** (1966), 360–363.