[6] E. HILLE, *Analytic Function Theory*, Vol. II, Blaisdell, Waltham, MA, 1962.
[7] A. ISERLES, *Order stars and a saturation theorem for first-order hyperbolics*, IMA J. Numer. Anal., 2 (1982), pp. 49–61.
[8] ———, *Order stars, approximations and finite differences* I. *The general theory of order stars*, this Journal, 16 (1985), pp. 559–576.
[9] A. ISERLES AND M. J. D. POWELL, *On the A-acceptability of rational approximations that interpolate the exponential function*, IMA J. Numer. Anal., 1 (1981), pp. 241–251.
[10] G. PICK, *Über die bescharankungen analytischer funktionen, welche durch vorgegebene funcktionswerte bewirkt werden*, Math. Ann., 77 (1916), pp. 7–237.
[11] G. WANNER, E. HAIRER AND S. P. NØRSETT, *Order stars and stability theorems*, BIT, 18 (1978), pp. 475–489.

SIAM J. MATH. ANAL.
Vol. 16, No. 4, July 1985

n-WIDTHS AND OPTIMAL INTERPOLATION OF TIME- AND BAND-LIMITED FUNCTIONS II*

AVRAHAM A. MELKMAN†

Abstract. Denote by $B(\sigma, T)$ the class of entire functions of exponential type σ which are bounded by 1 on the real axis outside $(-T, T)$. It is shown that this class, considered as a subset of $C[-T, T]$, has approximate dimension $2\sigma T/\pi$ in analogy to the Landau–Pollak–Slepian dimension theorem. More generally, the optimal subspaces and corresponding worst functions for the n-widths of $B(\sigma, T)$ are characterized. Prominently featured is the fact that it is possible to achieve the n-widths via interpolation, provided the sampling points are adroitly chosen. However, the interpolating functions differ from the standard ones.

1. Introduction. Denote by $B(\sigma, T)$ the class of entire functions of exponential type σ which are bounded by 1 on $(-\infty, -T)\cup(T, \infty)$. This paper mainly concerns the following problems.

(a) Given $\{t_i\}_1^n$ with $-T \leqq t_i \leqq T$, find an algorithm A^*: $C^n \to C[-T, T]$ which estimates $f \in B(\sigma, T)$ from the data $y = \{f(t_i)\}_1^n$ optimally in the sense of Micchelli and Rivlin [10], i.e. it achieves

$$E(t_1, \cdots, t_n) = \inf_A \max_{f \in B(\sigma, T)} \|f - Ay\|_T,$$

where $\|\cdot\|_T$ is the max norm on $[-T, T]$, and A is any map $C^n \to C[-T, T]$.

(b) Find the n-widths of $B(\sigma, T)$ with respect to $C[-T, T]$,

$$d_n(\sigma, T) = d_n(B(\sigma, T), C[-T, T])$$
$$= \min_{X_n \subset C[-T, T]} \max_{f \in B(\sigma, T)} \min_{y \in X_n} \|f - y\|_T,$$

where X_n is an n-dimensional subspace.

Clearly then $E(t_1, \cdots, t_n) \geqq d_n(\sigma, T)$; we will show that equality is achieved with the optimal choice of sampling points in (a). Of particular interest is $N(\sigma, T)$, the least n for which the n-width is 1 or less. This $N(\sigma, T)$ may be regarded as the "approximate dimension" of $B(\sigma, T)$, though its dimension is of course infinite. This point of view is best explained within the context of the original problem in which it arose. Regard $\varepsilon B(\sigma, T)$, $\varepsilon > 0$, as the set of those functions which are band-limited to (i.e. with Fourier transform supported on) $(-\sigma, \sigma)$ and simultaneously time-limited to $[-T, T]$ to within measurement accuracy ε; after all, outside $(-T, T)$ any $f \in \varepsilon B(\sigma, T)$ is pointwise indistinguishable from 0 within accuracy ε. Thus it is reasonable to define the approximate dimension of $\varepsilon B(\sigma, T)$ to be the dimension of the smallest subspace which contains for each $f \in \varepsilon B(\sigma, T)$ an element pointwise indistinguishable from f in $[-T, T]$ within accuracy ε.

The dimensionality problem is therefore the L_∞ version of the L_2 problem raised and dealt with by Landau and Pollak [6] and Slepian [12], of which we gave an account in a previous paper [9]. Analogously to the results found there we prove that the

*Received by the editors February 3, 1983, and in revised form February 27, 1984. This research was supported in part by the U.S. Army through its European office under contract DAJA 37-81-C-0234.
†Department of Mathematics, Ben-Gurion University, Beer-sheva, Israel.

approximate dimension of $B(\sigma,T)$ is $2\sigma T/\pi$ and that the approximation may proceed by interpolation.

Logan [7] analyzed the $n=0$ case, which is to find the function maximally concentrated in $[-T,T]$; his methods stimulated ours. Another source of inspiration has been the work of Boas and Schaeffer [3] (see also Ahiezer [1]). In fact, §2 is mostly devoted to showing how their arguments may be modified to solve the more general problem of characterizing the extremal f_0 for $\max(Lf: f\in B(\sigma,T))$, with L a linear functional. This characterization is of use for problem (a) when taking L of the form

$$Lf=f(t)-\sum_{i=1}^{n}\alpha_i f(t_i),$$

thereby obtaining the error in the pointwise estimation of $f(t)$ from the data $y=\{f(t_i)\}_1^n$ by means of the linear algorithm $Ay=\sum_{i=1}^{n}\alpha_i f(t_i)$. Moreover, Micchelli and Rivlin [10] have shown that for pointwise optimal estimation there always is such a linear algorithm which is optimal and

$$\min_{\{\alpha_i\}_1^n}\ \max_{f\in B(\sigma,T)}\left|f(t)-\sum_{i=1}\alpha_i f(t_i)\right|=\max_{\substack{f\in B(\sigma,T)\\ f(t_i)=0,\,i=1,\cdots,n}}|f(t)|.$$

Section 3 uses this observation to analyze the algorithm for pointwise optimal estimation and shows that it is also a globally optimal one.

In §4 it is shown that this linear optimal estimation based on the best set of interpolation points actually yields the n-widths, i.e.

$$d_n(\sigma,T)=\min_{\{t_i\}_1^n}\min_{A}\max_{f\in B(\sigma,T)}\|f-Ay\|_T,$$

a result similar to Micchelli, Rivlin and Winograd [11]. Finally in §5 the dimensionality of $B(\sigma,T)$ is calculated.

2. The maximum of a linear functional. This section summarizes the main results of Boas and Schaeffer [3] and the slight variations pertinent to the present setting.

Like them we are interested in the linear functional

$$Lf=\sum_{i=1}^{m}\sum_{j=0}^{n_i}a_i^{(j)}f^{(j)}(x_i) \tag{2.1}$$

with $a_i^{(j)}$ given real numbers, and x_i real points, and denote $l=\sum_{i=1}^{m}(n_i+1)$. The only differences are:

(i) we want to maximize Lf over the class $B_R(\sigma,T)$ of entire functions of exponential type σ which are real on the real axis and bounded by 1 on $(-\infty,-T)$ and (T,∞) instead of their $B_R(\sigma,0)$;

(ii) we require $x_i\in[-T,T]$, $i=1,\cdots,m$.

In the following theorem we gather all the information needed in later sections.

THEOREM 2.1. *The element f of $B_R(\sigma,T)$ for which*

$$Lf=\sup Lg \tag{2.2}$$

is unique and either constant or of type σ exactly. If not constant it satisfies a differential equation

$$\frac{\{f'(z)\}^2}{1-\{f(z)\}^2}=\frac{\sigma^2\{p(z)\}^2}{q(z)}$$

and is therefore of the form

$$f(z)=\sin\psi(z),\qquad \psi(z)=\sigma\int_0^z\{q(w)\}^{-1/2}p(w)\,dw+\sin^{-1}f(0).$$

Here $p(z)$ and $q(z)$ are monic polynomials with real coefficients.

Denote by $\cdots<\lambda_{-2}<\lambda_{-1}\leq -T$, $T\leq\lambda_1<\lambda_2<\cdots$ the points in $(-\infty,-T]$ and $[T,\infty)$ at which $|f(\lambda_i)|=1$. If $\lambda_{\pm1}=\pm T$ and $f'(\pm T)\neq 0$ set $s_{\pm}(z)=z\mp T$, otherwise $s_{\pm}(z)=1$. Let

$$s(z)=s_+(z)s_-(z),\qquad \nu=\text{degree } s. \tag{2.3}$$

(1) *The degree of p is at most $l-2+\nu$ and the zeros of $p(z)$ are precisely those zeros of f' different from the λ_i. Thus f nearly equioscillates outside $(-T,T)$ in the sense that $f(\lambda_i)f(\lambda_{i+1})=-1$ for all $i\leq -2$, $i\geq 1$ with the exception of at most $l-2+\nu$ values.*

(2) *f vanishes simply between successive λ_i such that $f(\lambda_i)f(\lambda_{i+1})=-1$ and has at most $l-1$ zeros in addition.*

Proof. We merely sketch the main points of the proof, referring to Boas and Schaeffer for details, whenever possible.

a. [3, Lemma 2.2]. sup $|Lg|$ is finite, positive and attained. Note: the proof of this in [3] needs only be supplemented by the fact, Logan [5], that if f is bounded by 1 on $(-\infty,-T)$ and (T,∞) then it is bounded by $\cosh\sigma T$ on $(-T,T)$.

b. [3, Lemma 3.3]. Let f be an extremal for (2.2). Then $|f(x)|=1$ for at least one x in $(-\infty,T]\cup[T,\infty)$; and if $g\in B_R(\sigma,T)$ satisfies $g(\lambda_i)=0$ $i=\pm1,\pm2,\cdots$ then $Lg=0$.

c. [3, Lemma 3.7]. If $g\in B_R(\sigma,T)$ satisfies $g(\lambda_i)=0$, $i=\pm1,\pm2,\cdots$ and $g(x)=O(|x|^{-(l-1)})$ as $|x|\to\infty$ then $g\equiv 0$. Note. Our requirement $x_i\in[-T,T]$, $i=1,\cdots,m$ removes the obstacle noted in [3], to completing their proof in case $n_i=0$, $i=1,\cdots,m$.

d. [3, p. 863]. In particular consider $f'(z)s(z)$ which vanishes at all λ_i. If it has r_1 additional zeros then let $p(z)$ be the monic polynomial with precisely these zeros. The function $f'(z)s(z)/p(z)$ is in $B_R(\sigma,T)$ and behaves as $O(|x|^{-(r_1-m)})$. Thus $r_1\leq l-2+m$.

e. [3, Lemma 3.2]. Similarly, if $(1-\{f(z)\}^2)s(z)$ has r_0 zeros in addition to the double ones at the λ_i then there is a function $g\in B_R(\sigma,T)$ vanishing at all λ_i and behaving like $O(|x|^{(r_0-m)/2})$. Thus $(r_0-m)/2<l-1$.

f. [3, p. 863]. Combining (d) and (e) shows that

$$\left\{\frac{f'(z)}{p(z)}\right\}^2\Big/\left\{\frac{1-f(z)^2}{q(z)}\right\}$$

is a zero-free entire function of exponential type bounded on the real axis and therefore constant.

g. In order to prove (2) let the additional zeros of f be comprised of α zeros in $[-T,T]$; 2β complex zeros which must come in pairs since $f(z)$ is real for real z; 2γ real zeros between successive λ_i, which must be even in number in order to preserve the sign change. Let $r(z)$ be the monic polynomial of degree $k=\alpha+2\beta+2\gamma$ with these zeros. Now if $k \geq l$ let h be a monic polynomial of degree k with all its zeros in $(-T,T)$ and such that $h^{(j)}(t_i)=0$ $j=0,\cdots,n_i$, $i=1,\cdots,m$. Set $d(z)=f(z)h(z)/r(z)$. Then sign $d(t)=\operatorname{sign} f(t)$ for $|t|\geq T$ while $d(t)\to f(t)$ as $|t|\to\infty$. Thus $|f(t)-\varepsilon d(t)|<1$, $|t|\geq T$, for small enough ε. However $L(f-\varepsilon d)=Lf$ contradicting the maximality of f.

3. Optimal estimation of a function from given data. Consider the optimal estimation of $f(t)$, $f\in B(\sigma,T)$, from its values $f(t_i)$, $i=1,\cdots,n$ with $-T\leq t_1\leq t_2\leq\cdots\leq t_n\leq T$ where, in case of coincident points, appropriate derivative evaluations should be taken. Since attention can be restricted to linear recoveries, e.g., Micchelli and Rivlin [10], the search is for optimal coefficients a_i^* and the extremal function f_0 achieving

$$(3.1)\qquad \left|f_0(t)-\sum_{i=1}^{n} a_i^* f_0(t_i)\right| = \min_{a_i\in C}\max_{g\in B(\sigma,T)}\left|g(t)-\sum_{i=1}^{n} a_i g(t_i)\right|.$$

As noted before, it is a result of Micchelli and Rivlin [10] that f_0 is at the same time the extremal for

$$(3.2)\qquad \max\left(|g(t)|: g\in B(\sigma,T),\ g(t_i)=0, i=1,\cdots,n\right).$$

Since f_0 may be assumed real on the real axis it is sufficient to consider only $g\in B_R(\sigma,T)$ in (3.1) and (3.2), and hence also only real a_i. Thus the knowledge that $f_0(t_i)=0, i=1,\cdots,n$ can be combined with the properties f_0 is endowed with as the extremal of a problem of type (2.2).

PROPOSITION 3.1. *Let f_0 be an extremal of problem (3.1). Then:*

(1) f_0 vanishes at the t_i while all its other zeros are real, simple and outside $[-T,T]$.

(2) f_0 equioscillates outside $(-T,T)$, i.e. denoting by

$$\cdots<\lambda_{-2}<\lambda_{-1}\leq -T,\quad T\leq\lambda_1<\lambda_2<\cdots$$

the points at which $|f_0(\lambda_i)|=1$ then $f_0(\lambda_i)=(-1)^{\rho+i+1}$ with $\rho=0$ for $i>0$ and $\rho=n$ for $i<0$ (assuming $f_0(\lambda_1)=1$).

(3) f_0' vanishes in $(-T,T)$ at precisely $n-1+\nu$ points μ_i, $i=1,\cdots,n-1+\nu$ separating the t_i, with $-T<\mu_1<t_1$ if $s_-(z)\not\equiv 1$, $t_n<\mu_{n-1+\nu}<T$ if $s_+(z)\not\equiv 1$.

Proof. Invoking Theorem 2.1 (2) with $l=n+1$ it follows that, in addition to the real, simple zeros between successive λ_i, f_0 vanishes only at the t_i. Therefore by Laguerre's theorem, Boas [2, 2.8], f_0' too has only real simple zeros separating those of f_0. Additional information on f_0' can be gleaned from Theorem 2.1 (1), to the effect that f_0' has in $(-T,T)$ at most ν zeros in addition to the $n-1$ zeros between t_i. If, for example, $s_-(z)\not\equiv 1$ meaning $f_0(-T)=1$, $f_0'(-T)>0$ (since $|f_0(t)|\leq 1$ for $t\leq -T$) then f_0' must vanish between $-T$ and t_1. A similar phenomenon occurs at T causing f_0' to possess precisely $n-1+\nu$ zeros in $(-T,T)$.

With notation as in Proposition 2.1 set

$$(3.3)\qquad h(z)=f_0'(z)s(z)\Big/\prod_{i=1}^{n-1+\nu}(z-\mu_i).$$

Thus $h(z)$ vanishes only at λ_i, $i=\pm1,\pm2,\cdots$

THEOREM 3.1. *The extremal f_0 of problem (3.2) is unique, up to a sign, and independent of t. The optimal estimate of $g(t)$, $|t|\leq T$, is effected by interpolating the data $g(t_i)$, $i=1,\cdots,n$ with a function in*

$$(3.4)\qquad \operatorname{span}\{h(t)t^i\}_0^{n-1}$$

... *mate errs at most by $|f_0(t)|$, with equality only for f_0.*

Proof. First we show that, with f_0 an extremal for (3.2), any function $g\in B(\sigma,T)$ has the representation

$$(3.5)\qquad g(z)=\sum_{i=1}^{n} g(t_i)\frac{h(z)}{h(t_i)}L_i(z)+\sum_{|k|=1}^{\infty} g(\lambda_k)\frac{h(z)\omega(z)}{h'(\lambda_k)\omega(\lambda_k)(t-\lambda_k)}$$

where $\omega(z)=\prod_{i=1}^{n}(z-t_i)$ and $L_i(z)$ are the Lagrange polynomials, $L_i(t_j)=\delta_{ij}$.

Indeed, let C_k be the rectangular contour consisting of $|y|=k$, $x=\eta_k\pm k$ where η_k is the point at which the maximum of f_0' in $(\lambda_k,\lambda_{k+1})$ is achieved. Then

$$I_k=\int_{C_k}\frac{f_0(z)}{h(z)\omega(z)}\frac{1}{z-t}dz$$

satisfies $\lim_{k\to\infty} I_k=0$ because, by Theorem 1, $h(z)$ behaves for large z like $\sin\sigma z/z^{n-1}$. Thus Cauchy's theorem yields the desired representation.

Now for $k\geq 1$ $\operatorname{sign} h'(\lambda_k)=\operatorname{sign} f_0''(\lambda_k)=-\operatorname{sign} f_0(\lambda_k)$ since λ_k is an extremum of f_0 (except for λ_1 if $f_0(T)=1$, but then $f_0'(T)<0$ leads to the same conclusion). Observe in addition that $\omega(\lambda_k)>0$ while $t-\lambda_k<0$ for $t\leq T$. Combined with similar results for $k\leq -1$ this yields

$$f_0(t)=h(t)\omega(t)\sum_{|k|\geq 1}\left|h'(\lambda_k)\omega(\lambda_k)(t-\lambda_k)\right|^{-1},\qquad t\leq T.$$

Thus for any $g\in B_R(\sigma,T)$ so that $|g(\lambda_k)|\leq 1$ and $t\in[-T,T]$

$$\left|g(t)-\sum_{i=1}^{n} g(t_i)\frac{h(t)}{h(t_i)}L_i(t)\right|\leq|f_0(t)|$$

as claimed.

To prove uniqueness use the representation (3.5) for any $g\in B_R(\sigma,T)$ vanishing at the t_i. We have

$$|g(t)|=\left|h(t)\omega(t)\sum_{|k|=1}^{\infty}\frac{g(\lambda_k)}{h'(\lambda_k)\omega(\lambda_k)(t-\lambda_k)}\right|\leq|f_0(t)|.$$

since $|g(\lambda_k)|\leq 1$. Moreover equality occurs for $|t|<T$ if and only if $g(\lambda_k)=\delta f_0(\lambda_k)$ with $|\delta|=1$ in which case formula (3.5) yields $g(t)=\delta f_0(t)$.

Up to now we have dealt with the estimate of a function at a point. Another approach is to attempt estimation of the function as a whole on $[-T,T]$. One looks then at algorithms $A: C^n\to C[-T,T]$, which map the data $y=(g(t_1),\cdots,g(t_n))$ to a complex valued continuous function, and the optimal one achieves

$$E(t_1,\cdots,t_n)=\min_{A}\max_{g\in B(\sigma,T)}\|g-Ay\|_T$$

where $\|g\|_T=\max(|g(t)|: -T\leq t<T)$.

However, the pointwise optimal estimate consisted of an interpolant which was in fact of exponential type σ and bounded on the real axis. Thus the following corollary is immediate.

COROLLARY. *The pointwise optimal estimate of $g(t)$ given in Theorem 3.1 is also the global optimal estimate of g on $[-T,T]$ and*

$$E(t_1,\cdots,t_n)=\|f_0\|_T.$$

4. Optimal sampling and n-widths. The previous section closed with a description of an optimal procedure for recovering a function from its values sampled at a given set of points. In this setting it is natural to ask for the optimal set of points at which to sample, i.e., those points which minimize $E(t_1,\cdots,t_n)$,

$$E(t_1^*,\cdots,t_n^*)=\min_{t_i} E(t_1,\cdots,t_n)$$
$$=\min_{t_i}\left(\max\|g\|_T: g\in B(\sigma,T),\ g(t_i)=0, i=1,\cdots,n\right).$$

The following theorem answers this question and at the same time provides a characterization of the n-widths of $B(\sigma,T)$. Before stating the theorem, let us briefly describe the notion of n-widths; for a fuller description consult Lorentz [8] or Pinkus [12].

The procedure of interpolating a function $g\in B(\sigma,T)$, at a fixed set of n points by the n functions (3.4) is one particular kind of linear approximation. It is conceivable, and indeed sometimes the case that a different approximation process from some n-dimensional subspace X_n of functions will yields a smaller worst case error. Thus one is led to the notion of the n-width (in the sense of Kolmogorov) defined as

$$d_n(\sigma,T)=d_n(B(\sigma,T);\ C[-T,T])=\min_{X_n\subset C[-T,T]}\ \max_{g\in B(\sigma,T)}\ \min_{h\in X_n}\|g-h\|_T.$$

The n-width is therefore the minimum possible worst case error incurred in approximating the set $B(\sigma,T)$ with a set of n functions. Of particular interest, of course, is the set of functions that achieves the n-width.

THEOREM 4.1. *There exists a unique function $F_n\in B_R(\sigma,T)$ with the following properties:*

(1) F_n equioscillates in $[-T,T]$ between the values $\pm\|F_n\|_T$ exactly $n+1$ times at the points $\rho_1<\cdots<\rho_{r+1}$, i.e. $F_n(\rho_i)=(-1)^{n+1-i}\|F_n\|$ $i=1,\cdots,n+1$.

(2) F_n equioscillates outside $(-T,T)$ between ± 1.

(3) If $\|F_n\|_T<1$ then $|F_n(\pm T)|=\|F_n\|_T$ and otherwise $|F_n(\pm T)|=1$.

(4) F_n has only the real simple zeros implied by (1) and (2). This function is the unique solution to the problem

$$\min_{t_i}\ \max_{\substack{g\in B(\sigma,T)\\ g(t_i)=0}}\|g\|_T. \tag{4.1}$$

The zeros t_i of F_n in $[-T,T]$ are an optimal sampling point set and

$$d_n(B(\sigma,T);\ C[-T,T])=E(t_1^*,\cdots,t_n^*)=\|F\|_T.$$

Furthermore the n-width is achieved through the approximation process of interpolation at the points t_i^* by the set of functions $\{h(t)t^j\}_0^{n-1}$, where

$$h(t)=\begin{cases}F_n'(t)(t^2-T^2)/\prod_{i=1}^{n+1}(t-\rho_i) & \text{if } \|F_n\|_T>1,\\ F_n'(T)/\prod_{i=2}^{n}(t-\rho_i) & \text{if } \|F_n\|_T\le 1.\end{cases}$$

Proof. For ease of reading the proof is divided up into several lemmas. First we prove the existence of F_n as claimed and then its uniqueness. This already solves the

optimal sampling problem. Since the concomittant interpolation procedure is a form of linear approximation it also follows that $d_n\le\|F_n\|_T$. The proof is then completed by showing $d_n\ge\|F_n\|_T$.

LEMMA 4.1. *There exists a function $F\in B_R(\sigma,T)$ with the properties claimed for it in Theorem 4.1 (1)–(4).*

Proof. We employ, almost verbatim, the method of Karlin and Studden [5, Thm. 10.1]. Given any $\xi=(\xi_0,\xi_1,\cdots,\xi_n)$ within the simplex

$$\xi_i\ge 0,\quad i=0,\cdots,n,\qquad \sum_{i=0}^{n}\xi_i=2T \tag{4.2}$$

construct interpolation points $t_i=-T+\sum_{k=0}^{i-1}\xi_k$, $i=1,\cdots,n$.

By Theorem 3.1 there exists a unique function f_ξ such that

$$\|f_\xi\|_T=\max(\|g\|_T: g\in B(\sigma,T),\ g(t_i)=0,\ i=1,\cdots,n) \tag{4.3}$$

with the normalization $f(T+\varepsilon)>0$ for small ε. From Theorem 2.1 f_ξ already has properties (2), (4). It remains to show that ξ can be chosen so that properties (1) and (3) hold.

With $t_0=-T$, $t_{n+1}=T$ let

$$\delta_i(\xi)=\max_{t_i\le t\le t_{i+1}}|f_\xi(t)|,\qquad i=0,\cdots,n$$

and, at the suggestion of A. Pinkus,

$$\delta(\xi)=\max_i\delta_i(\xi),$$
$$e_i(\xi)=\delta(\xi)-\delta_i(\xi),\quad i=0,\cdots,n,\qquad e_{n+1}(\xi)=e_0(\xi).$$

Note that $\delta_i(\xi)=0$ if and only if $\xi_i=0$, i.e. $t_i=t_{i+1}$, and that $e_i(\xi)\ge 0$ with equality for at least one i. Now, any f_ξ can equioscillate at most $n+1$ times in $[-T,T]$, by Proposition 3.1 (3). Thus the existence of F is equivalent to the existence of a ξ such that $\sum_{k=0}^{n}e_k(\xi)=0$.

Suppose to the contrary that

$$e(\xi)=\sum_{k=0}^{n}e_k(\xi)>0\quad\text{for all }\xi.$$

Then the mapping $\xi\to\xi'$ given by

$$\xi_i'=\frac{2Te_i(\xi)}{e(\xi)},\qquad i=0,\cdots,n \tag{4.4}$$

is well defined on the whole simplex (4.2). Moreover the mapping is continuous because, roughly speaking, a sequence of extremals converges to an extremal. More precisely, suppose ξ^k converge to ξ, and let $\{t_i^k\}$, $\{t_i\}$ be the corresponding points. Denote the extremal of (4.3) for ξ^k by f_k. Then f_k converges to f because $B(\sigma,T)$ is a normal family. Our claim is that $f=f_\xi$. Because of the uniqueness of f_ξ it suffices to prove that if $g\in B(\sigma,T)$, and $g(t_i)=0$ $i-1,\cdots,n$, then $|g(t)|\le|f(t)|$. Indeed, given any $\delta>0$ choose k large enough so that

$$|t_i-t_i^k|\le\delta\min_j|t-t_j|,\quad i=1,\cdots,n,\qquad |f_k(t)|\le(1+\delta)|f(t)|.$$

Consider then

$$g_k(t)=g(t)\prod_{i=1}^{n}\left(\frac{t-t_i^k}{t-t_i}\right).$$

Since $g_k(t_i^k)=0,\ i=1,\cdots,k$. It follows that $|g_k(t)|\leq|f_k(t)|$. But

$$|g_k(t)|\geq|g(t)|\prod_{i=1}^{n}\left(1-\left|\frac{t_i-t_i^k}{t-t_i}\right|\right)\geq|g(t)|(1-\delta)^n$$

implying $|g(t)|\leq|f(t)|(1+\delta)(1-\delta)^{-n}$ for any δ and therefore $|g(t)|\leq|f(t)|$.

Thus the mapping (4.4) has a fixed point ξ^*

$$\xi_i^*=\frac{2Te_i(\xi^*)}{e(\xi^*)}.$$

However $e_i(\xi^*)=0$ for some i, hence for that i, $\xi_i^*=0$ and therefore $\delta_i(\xi^*)=0$. But also $\delta(\xi^*)-\delta_i(\xi^*)=0$ and so $\delta(\xi^*)=0$, a contradiction.

LEMMA 4.2. *Let f_n be the extremal of problem* (3.2),

$$\max\left(|g(t)|: g\in B(\sigma,T),\ g(t_i)=0, i=1,\cdots,n\right).$$

If $f\in B(\sigma,T)$ oscillates $n+1$ times in $[-T,T]$, i.e. there exist extrema $-T\leq e_1<e_2<\cdots<e_{n+1}\leq T$ such that $(-1)^{n+1-i}f(e_i)>0$, then either $\min_i|f(e_i)|<\|f_n\|_T$ or f is a constant multiple of f_n.

Proof. Assume to the contrary that $\min_i|f(e_i)|\geq\|f_n\|_T$. Using the representation formula (3.5), based on f_n,

$$|f(t)-h(t)p(t)|\leq|f_n(t)|,\qquad -T\leq t\leq T,$$

where $h(t)p(t)$ is the interpolant to f based on the points t_i, from the set $\{h(t)t^k\}_0^{n-1}$. In particular

$$|f(e_i)-h(e_i)p(e_i)|\leq\|f_n\|_T$$

and therefore

$$\operatorname{sign}h(e_i)p(e_i)=\operatorname{sign}f(e_i)=(-1)^{n+1-i}.$$

If now $h(e_i)>0$, $i=1,\cdots,n+1$, as has to be the case if $\min|f(e_i)|>\|f_n\|_T$, then $(-1)^{n+1-i}p(e_i)\geq0$ which implies $p\equiv0$ because p is a polynomial of degree $n-1$. But $h(t)p(t)$ coincides with f at the t_i, i.e. $f(t_i)=0$, $i=1,\cdots,n$, and therefore by Theorem 3.1 $|f(t)|<|f_n(t)|$, $-T\leq t\leq T$, a contradiction.

Note that in any case $h(t)>0$ for $-T<t<T$. Thus a modification of the previous argument may be required only for $e_{n+1}=T$ with $f(T)=f_n(T)=\|f_n\|_T$ (or the analogous case $e_1=-T$). If $f_n(T)<1$ then from (3.3) $h(T)>0$ and no modification is called for. It remains therefore to investigate $f_n(T)=\|f_n\|_T=1$. In this case $h(t)=f_n'(t)/\prod_{i=1}^{n-1}(t-\mu_i)$ so that $h'(T)\neq0$. Differentiating (3.5) and evaluating at T yields

$$f'(T)=0=h'(T)p(T)+R_0+h'(T)R_1,$$

$$f_n'(T)=0=R_0+h'(T)S_1.$$

where the $k=1$ term has been split off as R_0. It is easily shown, by the same means employed in Theorem 3.1, that $|R_1|\leq S_1$. Hence

$$h'(T)p(T)=h'(T)(S_1-R_1)$$

proves that again $p(T)\geq0$.

LEMMA 4.3. *The function F_n obtained in Lemma* 4.1 *is the unique solution to problem* (4.1).

Proof. Let $s_i, i=1,\cdots,n$ be the zeros of F_n in $(-T,T)$. Then the method of Theorem 3.1 shows that there is a representation formula of the form (3.5) based on F_n and s_i. In particular it follows that F_n is the extremal of problem (3.2), i.e. if $g\in B(\sigma,T)$ satisfies $g(s_i)=0$, $i=1,\cdots,n$ then $|g(t)|\leq|F_n(t)|$, $-T\leq t\leq T$. Thus F_n is a candidate for the solution of (4.1).

On the other hand, if f_n is any other candidate, then the previous lemma shows

$$\min_i|F_n(\rho_i)|=\|F_n\|_T<\|f_n\|_T.$$

LEMMA 4.4. $d_n(\sigma,T)\geq\|F_n\|_T$.

Proof. We rely on Lorentz [8, Lemma 9.1] to the effect that if for each choice of complex signs $\sigma_i, i=1,\cdots,n+1$ there is a function $g\in B(\sigma,T)$ such that $g(\rho_i)=\sigma_i\|F\|_T$ then $d_n\geq\|F\|_T$.

Let $t_i, i=1,\cdots,n$ be the n zeros of F in $[-T,T]$ and denote $\omega(t)=\prod_{i=1}^{n}(t-t_i)$. Let $p(t)$ be the polynomial of degree n such that

$$p(\rho_i)=\sigma_i(-1)^{n+1-i}\omega(\rho_i),\qquad i=1,\cdots,n+1.$$

Write p in Lagrange form (thanks to T. Rivlin for pointing out this tack)

$$p(t)=\sum_{i=1}^{n+1}\sigma_i(-1)^{n+1-i}\omega(\rho_i)L_i(t).$$

Note that for $t>T$, $\operatorname{sign}(\omega(t)L_i(t))=\operatorname{sign}\omega(\rho_i)=(-1)^{n+1-i}$. Thus

$$|p(t)|\leq\sum_{i=1}^{n+1}|\omega(\rho_i)L_i(t)|$$

$$=(\operatorname{sign}\omega(t))\sum_{i=1}^{n+1}\omega(\rho_i)L_i(t)=|\omega(t)|,\qquad t\geq T.$$

Therefore the function

$$g(t)=\frac{F_n(t)p(t)}{\omega(t)}$$

satisfies

$$g(\rho_i)=\sigma_i(-1)^{n+1-i}F_n(\rho_i)=\sigma_i\|F_n\|_T,$$

$$|g(t)|\leq|F_n(t)|\leq1\quad\text{for } |t|\geq T,$$

and hence $g\in B(\sigma,T)$.

5. The dimensionality of time- and band-limited functions. For some small ε consider the set $\varepsilon B(\sigma,T)$. Any function in this set is of exponential type σ or less, and may therefore be regarded as band-limited to $(-\sigma,\sigma)$. Moreover such a function is almost time-limited to $(-T,T)$ because outside this interval it is indistinguishable from 0 to within accuracy ε.

Define the dimension, $N(\sigma,T)$, of this set to be the least n such that $d_n(\varepsilon B(\sigma,T)) = \varepsilon d_n(\sigma,T) \le \varepsilon$. $N(\sigma,T)$ is therefore the dimension of the smallest subspace of $C[-T,T]$ which contains for each $f \in B(\sigma,T)$ an element approximating f in the max norm on $[-T,T]$ to within ε. Such an approximant is indistinguishable from f at the same level of accuracy as f is known to be time-limited.

This notion of the dimension of the class of almost time- and band-limited functions was introduced and investigated in the $\mathscr{L}_2$ setting by Landau and Pollak [6] and Slepian [13]. A slightly improved version of their results is contained in [9]. The main result, that $N(\sigma,T)=2\sigma T/\pi$, is replicated here. We will see however that the corresponding natural basis is somewhat different. But first a preparatory proposition.

PROPOSITION 5.1. a. $d_{n+1}(\sigma,T)<d_n(\sigma,T)$.

b. *If* $T_1<T_2$ *then* $d_n(\sigma,T_1)<d_n(\sigma,T_2)$.

Proof.

a. Let F_n, F_{n+1} be the extremals corresponding to d_n, d_{n+1}, e.g. $d_n=\|F_n\|_T$. Recall that F_n (F_{n+1}) equioscillates precisely $n+1$ $(n+2)$ times. By Lemma 4.2 $d_{n+1}(\sigma,T)=\|F_{n+1}\|_T<\|F_n\|_T=d_n(\sigma,T)$.

b. Let F_n, G_n be the extremals corresponding to $d_n(\sigma,T_1)$, $d_n(\sigma,T_2)$. Note that $T_1<T_2$ implies $B(\sigma,T_1)\subseteq B(\sigma,T_2)$ and therefore $F_n\in B(\sigma,T_2)$. Applying Lemma 4.2 yields $d_n(\sigma,T_1)=\|F_n\|_T<\|G_n\|_T=d_n(\sigma,T_2)$ unless $F_n=G_n$. The latter however is impossible: if $\|G_n\|_T>1$ then $G_n(t)>1\ge F_n(t)$ for $T_1<t<T_2$; if $\|G_n\|_T\le 1$ then $\pm T_2$ must be two of its equioscillation points while F_n has all of its equioscillation points in $[-T_1,T_1]$.

THEOREM 5.1. *Let* $N(\sigma,T)$ *be the least* n *such that* $d_n(\sigma,T)\le 1$. *Then* $N(\sigma,T)=\lceil 2\sigma T/\pi\rceil=$ *the least integer not less than* $2\sigma T/\pi$. *In case* $2\sigma T/\pi$ *is an integer an optimal N-dimensional approximating subspace is spanned by the functions*

$$(5.1)\qquad \sin\sigma t \text{ and } \frac{\sin(\sigma t-k\pi)}{\sigma t-k\pi},\quad k=0,\pm1,\cdots,\pm(m-1)\qquad \text{if } \frac{2\sigma T}{\pi}=2m,$$

$$(5.2)\qquad \cos\sigma t \text{ and } \frac{\cos(\sigma t-r\pi)}{\sigma t-r\pi},\quad r=\pm\frac{1}{2},\frac{3}{2},\cdots,\pm\left(m-\frac{3}{2}\right)\qquad \text{if } \frac{2\sigma T}{\pi}=2m-1.$$

Proof. Observe that if $2\sigma T/\pi=2m$ then $F_{2m}(t)=\cos\sigma t$ has properties (1)–(4) of Theorem 4.1 with $\|F_{2m}\|_T=1$ and therefore $d_{2m}(\sigma,m\pi/\sigma)=1$. From Proposition 5.1 $d_{2m-1}(\sigma,m\pi/\sigma)>1$ and hence $N(\sigma,m\pi/\sigma)=2m$. Theorem 4.1 also yields the result that an optimal $2m$-dimensional subspace is spanned by

$$(t^j\sin\sigma t)\Big/\prod_{i=-(k-1)}^{k-1}(\sigma t-i\pi)\qquad j=0,\cdots,2k-1$$

which is equivalent to (5.1). Note that the approximation may proceed by interpolation at the points $r\pi/\sigma$, $r=\pm\frac{1}{2},\cdots,\pm(m-\frac{1}{2})$. The case $2\sigma T/\pi=2m-1$ is proved analogously. If now $2\sigma T/\pi\le 2m+1$ then from Proposition 5.1

$$1=d_{2m}(\sigma,m\pi/\sigma)<d_{2m}(\sigma,T),$$

$$d_{2m+1}(\sigma,T)<d_{2m+1}\left(\sigma,\left(m+\frac{1}{2}\right)\pi/\sigma\right)=1,$$

proving $N(\sigma,T)=2m+1=\lceil 2\sigma T/\pi\rceil$. The proof of the case $2m-1<2\sigma T/\pi\le 2m$ proceeds similarly.

Finally, let us mention those explicit n-widths that are easily derived:

1. $d_0(\sigma,T)=\cosh\sigma T$, extremal: $F_0(t)=\cos\sigma\sqrt{t^2-T^2}$;
2. for $0<\sigma T\le\pi/2$, $d_1(\sigma,T)=\sin\sigma T$, extremal: $F_1(t)=\sin\sigma t$;
3. for $0<\sigma T\le\pi$ $d_2(\sigma,T)=\sin((\sigma T)^2/2\pi)$, extremal: $F_2(t)=\cos\sigma\sqrt{t^2+\alpha^2}$ with $\sigma\alpha=\pi/2-(\sigma T)^2/2\pi$; an optimal subspace is spanned by

$$\frac{\sin\sigma\sqrt{t^2+\alpha^2}}{\sigma\sqrt{t^2+\alpha^2}}t^j,\qquad j=0,1$$

with interpolation at the points $\pm(T/2)\sqrt{2-(\sigma T/\pi)^2}$.

Furthermore, as noted by Jagerman [4], an easy asymptotic estimate is obtained via polynomial interpolation, namely

$$d_n(\sigma,T)\le \text{const}\,\frac{1}{n!}\left(\frac{\sigma T}{2}\right)^n.$$

A better estimate for small $n>2\sigma T/\pi$ is $c_1\exp[c_2(2\sigma T/\pi-n)]$, c_1 and c_2 constants. This estimate can be deduced from H. J. Landau's results as described in his recent manuscript *Extrapolating a band-limited function from its samples taken in a finite interval.*

Acknowledgments. This work would not have come to fruition without the stimulating conversations with Professors R. Boas and S. Fisher and the congenial hospitality of Northwestern University.

REFERENCES

[1] N. I. AHIEZER, *Extremal properties of entire functions of exponential type*, Amer. Math. Soc. Transl., 86 (1970), pp. 1–30.

[2] R. P. BOAS, *Entire Functions*, Academic Press, New York, 1954.

[3] R. P. BOAS AND A. C. SCHAEFFER, *Variational methods in entire functions*, Amer. J. Math., 79 (1957), pp. 857–884.

[4] D. JAGERMAN, *Information theory and approximation of band-limited functions*, Bell System Tech. J., 49 (1970), pp. 1911–1941.

[5] S. KARLIN AND W. J. STUDDEN, *Tchebycheff Systems: With Applications in Analysis and Statistics*, John Wiley, New York, 1966.

[6] H. J. LANDAU AND H. O. POLLAK, *Prolate spheroidal wave functions, Fourier analysis and uncertainty*. III, Bell System Tech. J., 41 (1962), pp. 1295–1336.

[7] B. LOGAN, JR., *Properties of high-pass signals*, Ph.D. dissertation, Columbia Univ., New York, 1965.

[8] G. G. LORENTZ, *Approximation of Functions*, Holt, Rinehart and Winston, New York, 1966.

[9] A. A. MELKMAN, *n-widths and optimal interpolation of time- and band-limited functions*, in Optimal Estimation in Approximation Theory, C. A. Micchelli and T. J. Rivlin, eds., Plenum Press, New York, 1977, pp. 55–68.

[10] C. A. MICCHELLI AND T. J. RIVLIN, *A survey of optimal recovery* in Optimal Estimation in Approximation Theory, C. A. Micchelli and T. J. Rivlin, eds., Plenum Press, New York, 1977, pp. 1–53.

[11] C. A. MICCHELLI, T. J. RIVLIN AND S. WINOGRAD, *The optimal recovery of smooth functions*, Numer. Math., 26 (1976), pp. 191–200.

[12] A. PINKUS, *n-Widths in Approximation Theory*, to be published by Springer-Verlag, New York.

[13] D. SLEPIAN, *On band width*, Proc. IEEE, 64 (1976), pp. 292–300.