JOURNAL OF APPROXIMATION THEORY **18**, 189–204 (1976)

On an Optimal Method for the Numerical Differentiation of Smooth Functions

C. A. MICCHELLI

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

Communicated by T. J. Rivlin

Received March 25, 1975

Given a smooth function f and a bound on its derivatives we construct an optimal linear estimator for $f'(0)$ which uses only function values of f.

INTRODUCTION

An important area of numerical analysis is concerned with the problem of numerical differentiation of real data. This problem, as is well known, is ill-conditioned since differentiation is an unbounded operator. There are, however, a number of methods for the computation of the derivative of a function using only function values (cf. [1, 3]). In regard to this problem, a natural theoretical question to consider is finding the minimum intrinsic error in such a computation and identifying if possible, a method for computing the derivative which achieves the minimum error. Such a question was posed by Newman in [10]. The purpose of this paper is solve Newman's problem, as well as a generalization of it.

Suppose we are given a function f in $W^n(R) = \{f : f^{(n-1)}$ abs. cont. on every finite interval, $f^{(n)} \in L^\infty(R)\}$ $(R = (-\infty, +\infty))$. We wish to compute $f'(0)$ using only function values of f. We will, however, allow for computational errors of magnitude $\leqslant \epsilon$ in the computed values of f. Given no further information on f we may, of course, encounter arbitrarily large errors. However, suppose we have some a priori bounds on the derivatives of f; then we may ask, with some hope of success, for the minimum error under this additional information. With these constraints in mind, let us now formulate the precise problem which we will be concerned with in this paper.

Given a set of real numbers $T = \{t_1, ..., t_{2m}\}$, not necessarily all distinct, we define the polynomial

$$p_{2m}(x) = \prod_{j=1}^{2m} (x - t_j) = a_0 + a_1 x + \cdots + a_{2m-1} x^{2m-1} + x^{2m}$$

and the associated constant coefficient differential operator

$$\mathscr{L}f = p_{2m}(D)f = a_0 f + a_1 f' + \cdots + a_{2m-1} f^{(2m-1)} + f^{(2m)}.$$

We assume throughout the paper that $T = -T$. Hence we may express the set T in the form $T = \{\pm t_j : j = 1, 2,..., m\}$, where $0 \leqslant t_1 \leqslant \cdots \leqslant t_m$. We require our f to belong to W^{2m} and to satisfy

$$\| \mathscr{L}f \| \leqslant \gamma, \tag{1}$$

where γ is some positive number and $\| \cdot \|$ denotes the sup-norm on R.

Let S be *any* mapping from $L^\infty(R)$ into R. Then we interpret $S(g)$ where $g \in L^\infty(R)$ and $\| f - g \| \leqslant \epsilon$ as an estimator for $f'(0)$. The error in this estimate for $f'(0)$ does not exceed

$$E_S = E_S(\gamma, \epsilon) = \sup_{\substack{\|\mathscr{L}f\| \leqslant \gamma \\ \|f-g\| \leqslant \epsilon}} | f'(0) - S(g)| \tag{2}$$

under constraint (1). The minimum error is defined to be

$$E = E(\epsilon, \gamma) = \inf_S E_S, \tag{3}$$

and $\hat{S}$ is called an optimal estimator for $f'(0)$, if $E = E_{\hat{S}}$.

Our purpose in this paper is to find E and $\hat{S}$. This was done in [10] for the case $T = \{0\}$ and $m = 1$. We will construct for any T and m an optimal estimator for $f'(0)$ which is a *linear* functional on $L^\infty(R)$. The value of the minimum error E will be identified and shown to be related to the Landau problem on R for the differential operator $\mathscr{L}f$. The Landau constant in the case $T = \{0\}$ was determined by Kolmogorov [5] and recently, from a different approach, by Cavaretta [2]. Our proofs rely on some results on cardinal interpolation contained in [8, 9] and follow the approach used by Schoenberg [11, 12], for some elementary cases of the Landau problem.

Section 1 contains the construction of an estimator for $f'(0)$ while Section 2 contains a proof that it is an optimal linear estimator.

1

We define the class of cardinal $\mathscr{L}$-splines by

$$\mathscr{S}_{2m-1}(T) = \{S : S \in C^{2m-2}(R), S\,|_{(\nu,\nu+1)} \in \pi_{2m-1}(T), \nu \in Z\}, \tag{4}$$

where

$$\pi_{2m-1}(T) = \{f : \mathscr{L}f = 0\}$$

and Z is the set of integers. An associated class of null splines is defined by

$$\mathscr{S}_{2m-1}^{0}(T) = \{S: S(\nu + \tfrac{1}{2}) = 0, \nu \in Z, S \in \mathscr{S}_{2m-1}(T)\}, \tag{5}$$

and an eigenspline is any function $S \in \mathscr{S}_{2m-1}^{0}(T)$ which satisfies the functional equation

$$S(x + 1) = \lambda S(x), \qquad x \in R, \tag{6}$$

for some real number λ. In [8], we proved that $\mathscr{S}_{2m-1}^{0}(T)$ is a subspace of dimension $2m - 1$ spanned by $2m - 1$ eigensplines $S_1(x),..., S_{2m-1}(x)$ which satisfy the equations

$$S_i(x + 1) = \mu_i S_i(x), \qquad i = 1, 2,..., 2m - 1, \tag{7}$$

for some constants $\mu_1 ,..., \mu_{2m-1}$ which satisfy the relations

$$\begin{aligned} \mu_1 < \mu_2 < \cdots < \mu_{2m-1} < 0,& \\ \mu_i \mu_{2m-i} = 1, \qquad i = 1, 2,..., m - 1,& \\ \mu_m = -1.& \end{aligned} \tag{8}$$

We may express $S_j(x)$ on [0, 1] in the following convenient way. Let $A_n(x; \lambda)$ denote the nth divided difference of the function $g(z) = e^{xz}(e^z - \lambda)^{-1}$ at the points $z = t_1 , t_2 ,..., t_{2m}$. When the zeros of $p_{2m}(x)$ are all distinct, $A_n(x; \lambda)$ has the simple form

$$A_{2m-1}(x; \lambda) = \sum_{j=1}^{2m} (1/p_{2m}'(t_j))(e^{xt_j}/(e^{t_j} - \lambda)). \tag{9}$$

The function $A_{2m-1}(\frac{1}{2}; \lambda)$ has exactly $2m - 1$ distinct negative zeros given by

$$A_{2m-1}(\tfrac{1}{2}; \mu_j) = 0, \qquad j = 1, 2,..., 2m - 1, \tag{10}$$

and we may express the eigensplines as

$$S_j(x) = A_{2m-1}(x; \mu_j), \qquad j = 1, 2,..., 2m - 1. \tag{11}$$

The eigenspline $S_j(x)$ has exactly one simple zero on [0, 1] occurring at $x = \frac{1}{2}$ and

$$\operatorname{sgn} S_j(x) = (-1)^{j-1}, \qquad j = 1, 2,..., 2m - 1, \tag{12}$$

for $x > \frac{1}{2}$. Also, the following symmetry relation holds.

$$S_{2m-j}(1 - x) = \mu_j S_j(x), \qquad j = 1, 2,..., 2m - 1. \tag{13}$$

The following differential operators are important in the study of cardinal $\mathcal{S}$-splines.

$$\begin{aligned} f^{[0]} &= f \\ f^{[1]} &= (D - t_1)f \\ f^{[2]} &= (D^2 - t_1^2)f \\ f^{[3]} &= (D - t_2)(D^2 - t_1^2)f \\ &\vdots \\ f^{[2m]} &= \prod_{j=1}^{m} (D^2 - t_j^2)f. \end{aligned} \tag{14}$$

In [9], we observed that the vectors

$$v_i = (S_i(1),..., S_i^{[2m-2]}(1)), \qquad i = 1, 2,..., 2m - 1,$$

are the eigenvectors of an oscillation matrix of order $2m - 1$. This fact is related to the functional equations (7). Thus we may conclude from the Gantmacker–Krein theorem, as in [9, Remark 2.3], that every nontrivial null spline

$$S(x) = \sum_{j=p}^{q} c_j S_j(x)$$

satisfies the inequality

$$p - 1 \leqslant S^-(S(1), S^{[1]}(1),..., S^{[2m-2]}(1)) \leqslant q - 1, \tag{15}$$

where $S^-(v)$ denotes the sign changes of the vector v where the zero components are discarded while $S^+(v)$ denotes the maximum number of sign changes of the vector v where the zero components may be replaced by $+1$ or -1.

Let $Z(y; (0, 1))$ denote the zeros (counting multiplicities) in $(0, 1)$ of a function $y \in \pi_{2m-1}(T)$. The following generalized version of the Bundan–Fourier lemma appears [9].

LEMMA 1. *If* $y \in \pi_{2m-1}(T)$ *and* $y^{[2m-1]}(1)\, y^{[2m-1]}(0) \neq 0$ *then*

$$\begin{aligned} Z(y; (0, 1)) \leqslant 2m - 1 - S^+(y(0), -y^{[1]}(0),..., (-1)^{2m-1} y^{[2m-1]}(0)) \\ - S^+(y(1), y^{[1]}(1),..., y^{[2m-1]}(1)). \end{aligned}$$

Remark. The above lemma holds on any finite interval and for any constant coefficient differential operator. There is also a version valid for any variable coefficient totally disconjugate differential operator (cf. [9]). Note that in the case of a constant coefficient differential operator we have a freedom in Lemma 1 in the ordering of the zeros of the characteristic polynomial of the differential operator. In (14) we chose the ordering to be $t_1, -t_1, t_2, -t_2, \ldots, t_m, -t_m$. This ordering of the zeros of $p_{2m}(x)$ is the most convenient choice for our purposes here.

Recently, Melkman [6, 7] presented an important extension of the Budan–Fourier lemma to spline functions and obtained as a result precise interpolation criteria for spline functions satisfying mixed boundary conditions.

We are now ready to prove our first lemma.

Let us assume that $m \geqslant 2$; the case $m = 1$ follows from the results in [10, 12] and will not be discussed here. This case does not require detailed properties of the eigensplines.

LEMMA 2. *There exists a unique function of the form*

$$K(x) = \sum_{j=m+1}^{2m-1} c_j S_j(x) \tag{16}$$

which satisfies the conditions

$$K^{(i)}(\tfrac{1}{2}) = 0, \qquad i = 2, 4, \ldots, 2(m-2), \tag{17}$$

$$K^{(2m-2)}(\tfrac{1}{2}) = \tfrac{1}{2}. \tag{18}$$

Proof. Suppose that $K(x)$ does not exist. Then there must exist a nontrivial function K of form (16) which satisfies the conditions $K^{(\nu)}(\frac{1}{2}) = 0$, $\nu = 2, 4, \ldots, 2m-2$. We now define

$$\begin{aligned} H(x) &= K(x), & x \geqslant \tfrac{1}{2}, \\ &= -K(1-x), & x \leqslant \tfrac{1}{2}. \end{aligned} \tag{19}$$

Since $K(\frac{1}{2}) = 0$ it follows from (7), (8), and (13) that $K \in \mathscr{S}^0_{2m-1}(R) \cap L^\infty(R)$. Thus $H(x)$ must be a multiple of $S_m(x)$. This is not possible unless $H(x) \equiv 0$. Thus the lemma is proved.

Remark. Relations (17) and (18) are equivalent to the equations

$$K^{[i]}(\tfrac{1}{2}) = 0, \qquad i = 2, \ldots, 2(m-2), \tag{20}$$

$$K^{[2m-2]}(\tfrac{1}{2}) = \tfrac{1}{2}. \tag{21}$$

LEMMA 3. *The function K constructed in Lemma 2 has the following properties.*

(i) $K(\nu + \frac{1}{2}) = 0$, $\nu \in Z$,

(ii) $K \in C^{2m-2}(R)$,

(iii) $\operatorname{sgn} K^{[2m-1]}(\nu^+) = -\operatorname{sgn} K^{[2m-1]}(\nu^-) = (-1)^{\nu-1}$, $\quad \nu = 1, 2, \ldots$,

(iv) $(-1)^{\nu+m} K(x) > 0$, $x \in (\nu + \frac{1}{2}, \nu + \frac{3}{2})$, $\quad \nu = 0, 1, 2, \ldots$.

Proof. Let us note that (i) and (ii) follow immediately from representation (16). From (7) and (15) we have the inequality

$$m \leqslant S^-(K^{[j]}(\nu))_0^{2m-2}, \qquad \nu = 1, 2,\ldots . \tag{22}$$

Using the relation $S^+((-1)^j c_j)_0^n + S^-(c_j)_0^n = n$, (22) has the equivalent form

$$S^+((-1)^j K^{[j]}(\nu))_0^{2m-2} \leqslant m - 2, \qquad \nu = 1, 2,\ldots . \tag{23}$$

We claim that $K^{[2m-1]}(\frac{1}{2}) K^{[2m-1]}(1^-) \neq 0$, otherwise, $K^{[2m-1]}(x) \equiv 0$ on $(\frac{1}{2}, 1)$. This fact with (17), (18), and (22) (when $\nu = 1$) results in a contradiction when we apply the Budan–Fourier lemma to K on $(\frac{1}{2}, 1)$ (relative to $\{f: f^{[2m-1]} = 0\}$). Thus we may again apply the Budan–Fourier lemma, now for $\pi_{2m-1}(T)$, and obtain

$$\begin{aligned} 0 \leqslant Z(K; (\tfrac{1}{2}, 1)) &\leqslant 2m - 1 - S^+((-1)^j K^{[j]}(\tfrac{1}{2}))_0^{2m-1} - S^+(K^{[j]}(1^-))_0^{2m-1} \\ &\leqslant 2m - 1 - (m - 1) - m = 0. \end{aligned}$$

Thus we conclude that

$$S^+((-1)^j K^{[j]}(\tfrac{1}{2}))_0^{2m-2} = S^+((-1)^j K^{[j]}(\tfrac{1}{2}))_0^{2m-1} = m - 1,$$

$$S^-(K^{[j]}(1))_0^{2m-2} = S^+(K^{[j]}(1^-))_0^{2m-1} = m,$$

and

$$Z(K; (\tfrac{1}{2}, 1)) = 0.$$

These facts with (17) and (18) easily imply that sgn $K'(\frac{1}{2}) = (-1)^{m-1}$, sgn $K(x) = (-1)^{m-1}$, $x \in (\frac{1}{2}, 1)$, and sgn $K^{[2m-2]}(1) =$ sgn $K^{[2m-1]}(1^-) = -1$. A similar deliberation with the Budan–Fourier lemma applied to the interval $(1, 2)$ gives us the equations

$$S^+((-1)^j K^{[j]}(1))_0^{2m-2} = S^+((-1)^j K^{[j]}(1^+))_0^{2m-1} = m - 2,$$

$$S^-(K^{[j]}(2))_0^{2m-2} = S^+(K^{[j]}(2^-))_0^{2m-1} = m,$$

and

$$Z(K; (1, 2)) = 1.$$

Thus we conclude that sgn $K(x) = (-1)^{m-1}$, $x \in (1, \frac{3}{2})$, sgn $K(x) = (-1)^m$, $x \in (\frac{3}{2}, 1)$, and sgn $K^{[2m-1]}(1^+) =$ sgn $K^{[2m-1]}(2^-) = 1$.

The remainder of the proof proceeds in the same manner. It may be formalized by induction on ν where proper account is taken of previously obtained sign information of $K(x)$ and Eqs. (22) and (23). We omit the obvious details.

Remark. Lemma 3 was proved in [11] for $T = \{0\}$ and $m = 2$.

Let us now extend the function $K(x)$, $x \geqslant \frac{1}{2}$, constructed in Lemma 2 to R as an odd function which we will denote by $H(x)$. $H(x)$ is defined by (19). This function will serve as the Peano kernel of a linear estimator for $f'(0)$. Before we turn to this matter we prove the following lemma which we will need later.

LEMMA 4. *There exists positive constants α and β such that, for any $f \in L^\infty(R) \cap W^{2m}(R)$,*

$$\| f^{(j)} \| \leqslant \alpha \| f \| + \beta \| \mathscr{L} f \|, \qquad j = 1, 2, \ldots, 2m-1. \tag{23}$$

Proof. Let $y(x) = \sum_{\nu=-m+1}^{m} f(\nu)\, \ell_\nu(x)$ be the unique function in $\pi_{2m-1}(T)$ which interpolates $f(x)$ at $-m+1, \ldots, m$. Then, according to Peano's kernel theorem,

$$L_j f = f^{(j)}(0) - \sum_{\nu=-m+1}^{m} f(\nu)\, \ell_\nu^{(j)}(0) = \int_{-m+1}^{m} R_j(t)\, \mathscr{L} f(t)\, dt,$$

where $R_j(t) = L_j(\phi((\cdot - t)_+))$ and $\phi((x-t)_+)$ is the Green's function of the initial value problem for the differential operator $\mathscr{L}$. Thus

$$| f^{(j)}(0) | \leqslant \alpha \| f \| + \beta \| \mathscr{L} f \|,$$

where $\alpha = \max_j \sum_\nu | l_\nu^{(j)}(0) |$ and $\beta = \max_j \int_{-m+1}^{m} | R_j(t) |\, dt$. Replacing $f(t)$ by $f(t+x)$ in the above inequality completes the proof of the lemma.

We define

$$c_\mu = H^{[2m-1]}(\mu^+) - H^{[2m-1]}(\mu^-) = H^{(2m-1)}(\mu^+) - H^{(2m-1)}(\mu^-); \tag{24}$$

then

LEMMA 5. *For any $f \in L^\infty(R) \cap W^{2m}(R)$*

$$f'(\tfrac{1}{2}) = \sum_{-\infty}^{+\infty} c_\mu f(\mu) - \int_{-\infty}^{+\infty} H(x)\, \mathscr{L} f(x)\, dx, \tag{25}$$

where

$$\operatorname{sgn} c_\mu = (-1)^{\mu-1}$$

and

$$\operatorname{sgn} H(x) = (-1)^{\nu+m-1}, \qquad x \in (\nu + \tfrac{1}{2}, \nu + \tfrac{3}{2}).$$

Proof. We integrate by parts

$$\begin{aligned}\int_{-N}^{+N} H(x)\, \mathscr{L} f(x)\, dx = \text{Boundary terms} \Big|_{-N}^{+N} &+ \sum_{-N}^{+N} c_\mu f(\mu) \\ &+ (H^{[2m-1]}(\tfrac{1}{2}^+) - H^{[2m-1]}(\tfrac{1}{2}^-))\, f(\tfrac{1}{2}) \\ &- (H^{[2m-2]}(\tfrac{1}{2}^+) - H^{[2m-2]}(\tfrac{1}{2}^-))\, f'(\tfrac{1}{2}).\end{aligned}$$

Since $H(x)$ decays exponentially fast at $\pm\infty$, Lemma 4 implies that as $N \to \infty$ the boundary terms converge to zero. This completes the proof of Lemma 5.

We now define the linear estimator for $f'(0)$ by

$$\hat{S}(f) = \sum_{-\infty}^{+\infty} c_\mu f(\mu - \tfrac{1}{2}).$$

The next part of the paper is devoted to proving that $\hat{S}f$ is an optimal estimator for $f'(0)$.

The key to the proof that $\hat{S}(f)$ is an optimal estimator for $f'(0)$ is the eigenspline $S_m(x)$ which was discarded in the construction of $\hat{S}f$. Let us now define the function

$$F(x) = \int_0^x S_m(t)\,dt \Big/ \int_0^{1/2} S_m(t)\,dt.$$

$F(x)$ inherits the following properties from $S_m(x)$.

$$\begin{gathered} F(x+1) = -F(x), \qquad F(1-x) = F(x), \\ \| F \| = F(\tfrac{1}{2}) = 1, \qquad F(\nu) = 0,\ \nu \in Z, \\ \operatorname{sgn} F(x) = F(\nu + \tfrac{1}{2}) = (-1)^\nu, \qquad x \in (\nu, \nu+1),\ \nu \in Z, \end{gathered}$$

and $F \in C^{2m-1}(R)$. Thus we see that $F^{[i]}(0) = F^{[i]}(1) = 0$, $i = 0, 2,\ldots, 2m-2$. Thus, applying the Budan–Fourier lemma to F on the interval $(0, 1)$ proves that $\operatorname{sgn} \mathscr{L}F(x) = (-1)^m$, $x \in (0, 1)$. Hence

$$\mathscr{L}F(x) = (-1)^{m+\nu} a, \qquad x \in (\nu, \nu+1),\ \nu \in Z,$$

where we define

$$a = \| \mathscr{L}F \| .$$

2. An Optimal Estimator for $f'(0)$

We introduce two constants. The first constant is the Landau constant for $\mathscr{L}f$ on R and is defined by

$$E_0(\epsilon, \gamma) = \sup_{\substack{\|f\| \leq \epsilon \\ \|\mathscr{L}f\| \leq \gamma}} | f'(0) |.$$

The second constant is the minimum error in estimating $f'(0)$ by *linear* estimators. Let $\alpha(t)$ be any function of bounded variation on R. Then we define

$$E_1(\epsilon, \gamma) = \min_{\alpha} \sup_{\substack{\|f-g\|\leqslant\epsilon \\ \|\mathscr{L}f\|\leqslant\gamma}} \left| f'(0) - \int_{-\infty}^{+\infty} g(t)\, d\alpha(t) \right|.$$

LEMMA 6. $\quad E_0(\epsilon, \gamma) \leqslant E(\epsilon, \gamma) \leqslant E_1(\epsilon, \gamma).$

Proof. The second inequality is obvious. The lower estimate is proved as follows. Let $\|f\| \leqslant \epsilon$ and $\|\mathscr{L}f\| \leqslant \gamma$. Then for *any* mapping $S : L^\infty(R) \to R$ we have

$$|f'(0) - S(0)| \leqslant E_S,$$
$$|f'(0) + S(0)| \leqslant E_S.$$

Thus by the triangle inequality $|f'(0)| \leqslant E_S$ and this completes the proof of the lemma.

THEOREM 1 (Landau Problem on R for $\mathscr{L}f$). *Define* $f_0(x) = \epsilon F(x)$. Then $E_0(\epsilon, \epsilon a) = f_0'(0)$. *Furthermore, if* $f'(0) = E_0(\epsilon, \epsilon a)$, *where* $\|f\| \leqslant \epsilon$, $\|\mathscr{L}f\| \leqslant \epsilon a$, *then* $f(x) = f_0(x)$, *for all* $x \in R$.

Proof. For any $f \in L^\infty(R) \cap W^n(R)$ we have the inequality

$$|f'(0)| \leqslant \left(\sum |c_\mu|\right) \|f\| + \left(\int_{-\infty}^{+\infty} |H(x)|\, dx\right) \|\mathscr{L}f\|. \tag{26}$$

This inequality follows from (25) where we have replaced $f(x)$ by $f(x - \frac{1}{2})$. Moreover, (25) also give us

$$f_0'(0) = \epsilon \sum c_\mu F(\mu - \tfrac{1}{2}) - \epsilon \int_{-\infty}^{+\infty} H(x + \tfrac{1}{2})\, \mathscr{L}F(x)\, dx.$$

Hence

$$E_0(\epsilon, \epsilon a) \leqslant \epsilon \sum_\mu |c_\mu| + \epsilon a \int_{-\infty}^{+\infty} |H(x)|\, dx.$$

Using the sign properties of $F(x)$ and $H(x)$ which we previously established we obtain

$$f_0'(0) = \epsilon \sum |c_\mu| + \epsilon a \int_{-\infty}^{+\infty} |H(x)|\, dx.$$

Thus we conclude that $E_0(\epsilon, \epsilon a) = f_0'(0)$. Moreover, suppose that f is any other function for which $f'(0) = E_0(\epsilon, \epsilon a)$ then

$$f'(0) = \epsilon \sum c_\mu f(\mu - \tfrac{1}{2}) - \epsilon \int_{-\infty}^{+\infty} H(x + \tfrac{1}{2})\, \mathscr{L}f(x)\, dx.$$

Therefore

$$0 = \sum | c_\mu | [\operatorname{sgn} c_\mu f(\mu - \tfrac{1}{2}) - 1]$$
$$+ \int_{-\infty}^{+\infty} | H(x + \tfrac{1}{2})| [-\operatorname{sgn} H(x + \tfrac{1}{2}) \mathscr{L}f(x) - a] \, dx,$$

and we conclude that

$$f(\mu + \tfrac{1}{2}) = (-1)^\mu = F(\mu + \tfrac{1}{2}),$$
$$\mathscr{L}f(x) = \mathscr{L}F(x) = (-1)^{m+\mu} a, \qquad x \in (\mu, \mu + 1).$$

Since $f, F \in C^{2m-1}(R)$ we conclude that $f - F \in \pi_{2m-1}(T)$. However, $f - F$ has and infinite number of zeros. This implies $f = F$ and the theorem is proved.

THEOREM 2 (Optimal Estimation of $f'(0)$).

$$E_0(\epsilon, \epsilon a) = E(\epsilon, \epsilon a) = E_1(\epsilon, \epsilon a) = E_S(\epsilon, \epsilon a).$$

Proof. In view of Lemma 6 and Theorem 1 it is sufficient to prove that $| f'(0) - \hat{S}(g)| \leqslant f_0'(0)$ for all f, g with $\| \mathscr{L}f \| \leqslant \epsilon a$, $\| f - g \| \leqslant \epsilon$.

$$| f'(0) - \hat{S}(g)| = \left| f'(0) - \sum c_\mu f(\mu - \tfrac{1}{2}) + \sum c_\mu (f(\mu - \tfrac{1}{2}) - g(\mu - \tfrac{1}{2})) \right|$$
$$\leqslant \epsilon a \int_{-\infty}^{+\infty} | H(x)| \, dx + \epsilon \sum | c_\mu |$$
$$= f_0'(0).$$

This inequality proves Theorem 2.

Theorem 2 solves our problem when $\epsilon a = \gamma$. To treat the general case we use the idea implicit in [5], and make a change of scale. To this end, we introduce the family of differential operators

$$\mathscr{L}_h f = \prod_{j=1}^{2m} (D - ht_j) f, \; h > 0.$$

We construct for $\mathscr{L}_h$, as previously done for $\mathscr{L}$, the functions $H_h(x)$ and $F_h(x)$. Now, we scale back to the operator $\mathscr{L}$ by introducing the functions $F(x; h) = F_h(x/h)$, $H(x; h) = h^{2m-2} H_h((x/h) + \tfrac{1}{2})$ and the linear functional

$$\hat{S}_h(f) = h^{-1} \sum_{-\infty}^{+\infty} c_\mu(h) \, f(\mu h - (h/2)),$$

where $c_\mu(h)$ is the jump in the $(2m-1)$st derivative of $H_h(x)$ at $x = \mu$. Thus we have

$$f'(0) = h^{-1} \sum_{-\infty}^{+\infty} c_\mu(h) f(\mu h - (h/2)) - \int_{-\infty}^{+\infty} H(x; h)\, \mathscr{L}f(x)\, dx, \qquad (27)$$

and

$$\begin{aligned} F(x+h; h) &= -F(x; h), \\ F(h-x; h) &= F(x; h), \\ F(h/2; h) &= \tfrac{1}{2}, \\ \mathscr{L}F(0; h) &= (-1)^m e(h), \end{aligned}$$

where

$$a(h) = \| \mathscr{L}F(\cdot\, ; h) \| .$$

Theorems 1 and 2 extend immediately to

THEOREM 3. *For any $h > 0$ and $m \geqslant 2$*

(a) $E_0(\epsilon, \epsilon a(h)) = f_0'(0)$, *where* $f_0(x) = \epsilon F(x; h)$. *If* $E_0(\epsilon, \epsilon a(h)) = f'(0)$ *for some* f, $\|f\| \leqslant \epsilon$ *and* $\|\mathscr{L}f\| \leqslant \epsilon a(h)$, *then* $f(x) = f_0(x)$.

(b) $E_0(\epsilon, \epsilon a(h)) = E(\epsilon, \epsilon a(h)) = E_1(\epsilon, \epsilon a(h)) = E_{\mathcal{S}_h}(\epsilon, \epsilon a(h))$.

We may now attempt to adjust the value of h so that $\epsilon a(h) = \gamma$ and solve our problem. This, however, is not always possible. In the next several lemmas we examine this question.

LEMMA 7. *$a(h)$ is a strictly monotonic function on* $(0, \infty)$.

Proof. Let us suppose that $a(h_1) = a(h_1)$ and $h_1 \neq h_2$. Then according to the uniqueness assertion of Theorem 3 we conclude that $F(x; h_1) = F(x; h_2)$ for all x. We assume without loss of generality that $h_1 < h_2$. Then

$$\begin{aligned} 0 &= F^{(2m)}(h_1^+; h_2) - F^{(2m)}(h_1^-; h_2) \\ &= \mathscr{L}F(h_1^+; h_1) - \mathscr{L}F(h_1^-; h_1) \\ &= 2(-1)^{m+1} a(h_1). \end{aligned}$$

This contradiction implies that $h_1 = h_2$. Thus $a(h)$ is a strictly monotone function and this completes the proof.

LEMMA 8.

(a) $\lim_{h\to 0} h^{-2m} a(h) = d = (-1)^m E_{2m}/(2m)!\, 2^{2m}$,

where E_n is the nth *Euler number.*

(b) $\lim_{h\to 0} h^{-2m} F(x; h) = d^{-1} F_0(x)$, *where F_0 is the unique function defined by*

$$F_0^{(j)}(0) = \delta_{2m,j}, \qquad j = 0, 1,..., 2m,$$
$$F(x) = -F(-x), \qquad x \in R,$$

and

$$D\mathscr{L}(D) F_0(x) = 0, \qquad x > 0.$$

Proof. Using the contour integral representation for divided differences we may express $F(x; h)$ in the form

$$F(x; h) = \frac{\oint_c (e^{xz}/z(e^{hz} + 1)\, p_{2m}(z))\, dz}{\oint_c (e^{hz/2}/z(e^{zh} + 1)\, p_{2m}(z))\, dz},$$

where c is any contour containing in its interior the zeros of $zp_{2m}(z)$ but not the poles of $(e^{zh} + 1)^{-1}$. Therefore

$$a(h) = (-1)^m \Big/ \frac{1}{2\pi i} \oint_c \frac{\operatorname{sech}(hz/2)}{zp_{2m}(z)}\, dz.$$

By using the power series expansion of sech z, (a) easily follows. The proof of (b) now is a consequence of the above expression for $F(x; h)$. This proves Lemma 8.

We define $s = \prod_{j=1}^m t_j^2$.

LEMMA 9.

(a) $\lim_{h\to\infty} F(x; h) = F_\infty(x)$;

when $s = 0$, then $F_\infty(x) = 0$, for all x. For $s > 0$, $F_\infty(x)$ is the unique function defined by conditions

$$F_\infty^{(j)}(0) = 0, \qquad j = 0, 2,..., 2m-2, \qquad F_\infty(\infty) = 1, \tag{28}$$

and

$$D \prod_{j=1}^{m} (D + t_j) F_\infty(x) = 0, \qquad x > 0. \tag{29}$$

(Recall that $0 \leqslant t_1 \leqslant \cdots \leqslant t_m$.) Also, F_∞ is an even function in $C^{2m-1}(R)$ which has only one knot at zero.

(b) *In all cases*, $\lim_{h\to\infty} a(h) = s$.

Proof. Let us first observe that F_∞ is uniquely determined by the above conditions. If there were two functions satisfying these conditions then the difference between them which we will call $G(x)$ satisfies the equation $\prod_{j=1}^m (D + t_j) G(x) = 0$, $x > 0$. The function $G(x) + G(-x)$ is in $\pi_{2m-1}(T)$ and has a $(2m-1)$st order zero at zero. Thus $G(x) = -G(-x)$ for all x. This is impossible unless $G(x) \equiv 0$. Now, the limit of any convergent subsequence of $F(x; h)$, which we denote by $\bar{F}$, must clearly satisfy (28). Furthermore, since $\| F(0, h)\| = F(h/2; h) = 1$, the limit must also be bounded. Hence it must satisfy the differential equation $D \prod_{j \in J} (D + t_j) \bar{F} = 0$, where $J = \{j : t_j > 0\}$. Therefore we may argue as before that if $s = 0$ then $\bar{F} \equiv 0$. The remainder of the proof of (a) follows easily. When $s = 0$, then (a) clearly implies that (b) holds. For $s > 0$, $F_\infty(x) = 1 + y(x)$, where $y \in \pi_{2m-1}(T)$. Thus $\lim_{h\to\infty} a(h) = (-1)^m \lim_{h\to\infty} \mathscr{L}F(0; h) = (-1)^m \mathscr{L}F_\infty(0) = s$, which verifies (b) and completes the proof.

Lemmas 8 and 9 imply that $a(h)$ is a strictly decreasing function on $(0, \infty)$ whose range is (s, ∞). If $s = 0$, then for any $\gamma > 0$ there exists a unique $h_\gamma > 0$ such that $\epsilon^{-1}\gamma = a(h_\gamma)$. Thus Theorem 3 gives us the value of $E(\epsilon, \gamma)$ and solves our problem.

When $s > 0$, then we may adjust h only for $\epsilon^{-1}\gamma$ in the range of $a(h)$, that is,

$$s\epsilon < \gamma.$$

When $s\epsilon \geqslant \gamma$ the nature of the extremal solution changes from a "perfect" $\mathscr{L}$-spline given by $\epsilon F(x; h)$ to the function F_∞ which has only *one knot* at zero. Our next lemma examines the form of the optimal differentiation formula (27) when $h \to \infty$ and $s > 0$.

We define the function

$$\Lambda(t) = (1/2\pi) \int_{-\infty}^{+\infty} (e^{itu}/p_{2m}(iu))\, du.$$

This function has the following properties.

$$\Lambda \in C^{2m-2}(R),$$
$$\Lambda(t) = \Lambda(-t),$$
$$\Lambda^{(2m-1)}(0+) = \tfrac{1}{2},$$

and

$$\prod_{j=1}^{m} (D + t_j)\, \Lambda(t) = 0, \qquad t > 0.$$

LEMMA 10. *When $s > 0$, $\lim_{h\to\infty} H(x; h) = \Lambda'(-x)$, uniformly for all x, and for any function $f \in W^n(R)$ with $f(x) = o(e^{t_1|x|})$, $x \to \pm\infty$,*

$$f'(0) = -\int_{-\infty}^{+\infty} \Lambda'(-t)\, \mathscr{L}f(t)\, dt.$$

Proof. According to the proof of Theorem 2 we have

$$a(h) \int_{-\infty}^{+\infty} | H(x; h)| \, dx \leqslant F'(0; h).$$

Thus Lemma 9 implies that $\int_{-\infty}^{+\infty} | H(x; h)| \, dx$ is uniformly bounded in h. The limit of any convergent subsequence $\bar{H}(x)$ must satisfy the conditions

$$\bar{H}^{(j)}(0) = 0, \qquad j = 0, 2, \ldots, 2(m-2), \qquad \bar{H}^{2m-2}(0^+) = \tfrac{1}{2},$$

$$\bar{H}(x) = -\bar{H}(-x), \qquad \prod_{j=1}^{m} (D + t_j) \, \bar{H}(x) = 0, \qquad x > 0.$$

These properties characterize the function $\Lambda'(-x)$. This completes the proof of the lemma.

Remark. For any $f \in W^n(R)$ *with* $f(x) = o(e^{t_1|x|})$, $x - \pm\infty$,

$$f(x) = \int_{-\infty}^{+\infty} \Lambda(t - x) \, \mathscr{L}f(t) \, dt$$

(cf. Hirshman and Widder [4]).

Finally, we summarize in our last theorem the complete solution of our problem.

THEOREM 4. *Given* $\epsilon > 0$, $\gamma > 0$, *and* $m \geqslant 2$.

(a) *If* $s\epsilon < \gamma$ *then there exists a unique positive number* h *such that* $a(h) = \epsilon^{-1}\gamma$ *and*

$$E(\epsilon, \gamma) = E_{S_h}(\epsilon, \gamma) = E_0(\epsilon, \gamma) = \epsilon F'(0; h).$$

Furthermore, if f *is any other function with* $E_0(\epsilon, \gamma) = f'(0)$, $\| f \| \leqslant \epsilon$, $\| \mathscr{L}f \| \leqslant \gamma$, *then* $f(x) = \epsilon F(x; h)$.

(b) *If* $s\epsilon \geqslant \gamma$, *then*

$$E(\epsilon, \gamma) = E_0(\epsilon, \gamma) = \epsilon F_\infty{}'(0),$$

and an optimal estimator for $f'(0)$ *is* $S(f) = 0$. *If* $E_0(\epsilon, \gamma) = f'(0)$, $\| f \| \leqslant \epsilon$, $\| \mathscr{L}f \| \leqslant \gamma$, *then* $f(x) = \epsilon F_\infty{}'(x)$.

Remark. An optimal estimator for $f'(t)$ may be easily obtained from our previous discussion by replacing $f(x)$ with $f(x + t)$. This device allows us to obtain an optimal estimator for the differentiation *operator* among all operators (possibly nonlinear) which map $L^\infty(R) \to L^\infty(R)$.

This paper is based on a lecture given at the Weizmann Institute of Science, Rehovot, Israel, in January 1974. The results presented here were substantially influenced by an earlier lecture, also given at the Weizmann Institute, by I. J. Schoenberg on the Landau problem for $D^2 \pm 1$. At the time that our results were obtained the article "Cardinal Interpolation and Spline Functions," Part VIII, which deals with the Landau problem for D^n, and promise in [11], had not appeared.[1]

The construction of the optimal estimator for $f'(0)$ was suggested by the discussion in [10] for the case $m = 1, T = \{0\}$, and in [11] for $m = 2$, $T = \{0\}$. The idea of using the Gantmacher–Krein theorem and the Budan–Fourier lemma for the general case discussed here is also used in [8] to obtain exact error estimates for cardinal $\mathscr{L}$-spline interpolation. The discussion in this paper (as well as in [8, Section 5]) easily extends to symmetric *odd* order differential operators $D \prod_{j=1}^{m} (D^2 - t_j^2)$. However, the treatment for an arbitrary nth order differential operator $\prod_{j=1}^{n} (D - t_j)$, $t_1 ,..., t_n$, real, requires further examination. In this regard, the recent paper, "Landau-Type Inequalities for Some Linear Differential Operators," by A. Sharma and J. Tzimbalario (preprint), solves the Landau problem for $\prod_{j=1}^{n} (D - t_j)$ by using a method employed in [2]. Although this method is elementary, using only Rolle's theorem and an approximation argument, it does not yield the *uniqueness* of the extremal function nor does it solve the *dual* version of the Landau problem. This latter fact is crucial to the point of view taken in this paper.

References

1. R. S. Anderssen and P. Bloomfield, On the numerical differentiation of data, Department of Statistics, Princeton University, Tech. Report No. 13, Ser. 2, 1972.
2. A. S. Cavaretta, Jr., An elementary proof of Kolmogorov's theorem, *Amer. Math. Monthly* **81** (1974), 480–486.
3. J. Cullum, Numerical differentitation and regularization, *SIAM J. Numer. Anal.* **8** (1971), 254–265.
4. I. I. Hirshman and D. V. Widder, "The Convolution Transform," Princeton Univ. Press, Princeton, N.J., 1955.
5. A. Kolmogorov, On inequalities between the upper bounds of successive derivatives of an arbitrary function on an infinite interval, *Amer. Math. Soc. Transl., Ser.* 1 **2** (1962), 233–243.

[1] Meanwhile the following paper appeared: Carl de Boor and I. J. Schoenberg, Cardinal Interpolation and Spline Functions. VIII. The Budan–Fourier Theorem For Splines and Applications, in Spline Functions Karlsruhe 1975 (K. Böhmer, G. Meinardus, W. Schempp, Eds.) Lecture Notes in Mathematics Vol. 501, Springer–Verlag, Berlin, 1976.

6. A. A. MELKMAN, The Budan–Fourier theorem for splines, *Israel J. Math.* **19** (1974), 256–263.
7. A. A. MELKMAN, Interpolation by splines satisfying mixed boundary conditions, *Israel J. Math.* **19** (1974), 369–381.
8. C. A. MICCHELLI, Cardinal $\mathscr{L}$-splines, *in* "Studies in Splines and Approximation Theory" (S. Karlin, C. A. Micchelli, A. Pinkus, and I. J. Schoenberg, Eds.), Academic Press, New York, 1976.
9. C. A. MICCHELLI, Oscillation matrices and cardinal spline interpolation, *in* "Studies in Splines and Approximation Theory" (S. Karlin, C. A. Micchelli, A. Pinkus, I. J. Schoenberg, Eds.), Academic Press, New York, 1976.
10. D. J. NEWMAN, Numerical differentiation of smooth data, preprint.
11. I. J. SCHOENBERG, The elementary cases of Landau's problem of inequalities between derivatives, *Amer. Math. Monthly*, **80** (1973), 121–158.
12. I. J. SCHOENBERG, Notes on spline functions. VI. Extremum problems of the Landau type for the differential operator $D^2 \pm 1$, *in* "Studies in Splines and Approximation Theory" (S. Karlin, C. A. Micchelli, A. Pinkus, and I. J. Schoenberg, Eds.). Academic Press, New York, 1976.