JOURNAL OF APPROXIMATION THEORY 54, 235–249 (1988)

Boundary Effects on Convergence Rates for Tikhonov Regularization*

FLORENCIO I. UTRERAS

Departamento de Matemáticas y Ciencias de la Computación, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Casilla 170, Correo 3, Santiago, Chile

Communicated by P. L. Butzer

Received April 27, 1984; revised April 25, 1986

We consider the Tikhonov regularizer f_λ of a smooth function $f \in H^{2m}[0, 1]$, defined as the solution (see [1]) to

$$\min_{u \in H^m[0,1]} \{\lambda |u|_m^2 + |f-u|_0^2\}, \qquad \lambda > 0.$$

We prove that if

$$f^{(j)}(0) = f^{(j)}(1) = 0, \qquad j = m, ..., k < 2m-1,$$

then

$$|f - f_\lambda|_j^2 \leqslant R\lambda^{(2k-2j+3)/2m}, \qquad j = 0, ..., m.$$

A detailed analysis is given of the effect of the boundary on convergence rates.

1. INTRODUCTION

Let $H^k[0, 1]$ be the Sobolev space

$$\{u\colon [0,1] \to \mathbb{R} \mid u, u', ..., u^{(k-1)} \text{ are absolutely continuous and } \int_0^1 (u^{(k)}(t))^2\, dt < +\infty\}$$

and let $|\cdot|_k$ denote the seminorm

$$|u|_k^2 = \int_0^1 (u^{(k)}(t))^2\, dt. \tag{1}$$

* Partially supported by Fondo Nacional de Ciencia y Tecnología under Grant 755/82.

0021-9045/88 $3.00

Also, for m a positive integer and $\lambda > 0$, let f_λ be the Tikhonov regularizer (cf. [12]) of f defined as the unique solution to the minimization problem

$$\min_{u \in H^m[0,1]} \{\lambda |u|_m^2 + |f-u|_0^2\}. \tag{2}$$

The existence and uniqueness of f_λ is well known (see, for example, [8]). The analysis of the properties of f_λ as λ goes to zero are of fundamental importance in the study of convergence rates of smoothing splines as has been established by Ragozin [8], Wahba [14], Craven and Wahba [3], Utreras [13], and others.

Ragozin [8] gave estimates for $|f-f_\lambda|_j$, $j=0, 1, ..., k$, with $f \in H^k[0,1]$, $k \leqslant m$. His main result is

THEOREM 1 (Ragozin). *For $j \leqslant k \leqslant m$ there exist constants $\beta = \beta(m, k, j)$ such that for $f \in H^k[0,1]$*

$$|f-f_\lambda|_j^2 \leqslant \beta \lambda^{(k-j)/m} |f|_k^2. \tag{3}$$

Thus for $j=k=m$ we have

$$|f-f_\lambda|_m^2 \leqslant \beta |f|_m^2. \tag{4}$$

Inequality (4) does not allow us to prove that $|f_\lambda - f|_m$ goes to zero as $\lambda \to 0$. Our aim in this paper is to give sharper estimates for the error $|f-f_\lambda|_j$. Moreover, we analyze how the values of f and its derivatives at 0 and 1 affect the convergence rates.

To do this, in Section 2 we write the solution f_λ as an expansion in terms of the eigenfunctions of the operator D^{2m} satisfying appropriate boundary conditions; we also give an expression for the error $|f-f_\lambda|_0^2$. In Section 3 we give a detailed discussion of the properties of the Fourier coefficients (or Birkhoff coefficients) of f in terms of the values of f and its derivatives at the end points of the interval. Finally, in the last section we apply this result to the study of convergence rates for the Tikhonov regularization procedure. We find that these bounds are strongly dependent upon the boundary conditions that f and its derivatives satisfy at 0 and 1.

2. THE FOURIER EXPANSION

Consider the eigenvalue problem

$$\begin{aligned} &(-1)^m D^{2m}\psi = \mu\psi, \\ &\psi^{(j)}(0) = \psi^{(j)}(1) = 0, \qquad j = m, ..., 2m-1. \end{aligned} \tag{5}$$

It is well known (cf. [6]) that the eigenvalues μ_i of (5) satisfy

$$\mu_0 = \mu_1 = \mu_2 = \cdots = \mu_{m-1} = 0,$$
$$\mu_i > 0, \qquad i \geqslant m.$$

Moreover, the eigenspace corresponding to the eigenvalue 0 is P_{m-1}, the set of all polynomials of degree $m-1$ or less.

Let $\psi_0, ..., \psi_{m-1}$ be an orthogonal basis for P_{m-1} such that

$$\int_0^1 \psi_i(x)\,\psi_j(x)\,dx = \begin{cases} 1, & i=j, \\ 0, & i \neq j, \end{cases} \tag{6}$$

and let $\{\psi_i\}$, $i \geqslant m$, be an orthonormal set of eigenfunctions of (5) where

$$\mu_m \leqslant \cdots \leqslant \mu_i \leqslant \mu_{i+1} \cdots$$

and each μ_i appears a number of times equal to its multiplicity. Then $\{\psi_i\}_0^\infty$ is a complete set of orthonormal functions in $H^0[0,1] = L^2[0,1]$ and $f \in H^k[0,1]$ can be expanded in the generalized Fourier series

$$f = \sum_{n \geqslant 0} \hat{f}_n \psi_n, \tag{7}$$

where

$$\hat{f}_n = \int_0^1 f(x)\,\psi_n(x)\,dx. \tag{8}$$

Let $u \in H^m[0,1]$. Then, for $n \geqslant m$

$$\begin{aligned} \hat{u}_n &= \int_0^1 u(x)\,\psi_n(x)\,dx \\ &= \frac{(-1)^m}{\mu_n} \int_0^1 u(x)\,\psi_n^{(2m)}(x)\,dx \\ &= \frac{1}{\mu_n} \int_0^1 u^{(m)}(x)\,\psi_n^{(m)}(x)\,dx. \end{aligned} \tag{9}$$

But $\psi_n^{(m)}$ is also a complete orthogonal system in $L^2[0,1]$ (see [5, p. 147]) and $u^{(m)} \in L^2[0,1]$. Thus

$$\sum_{n \geqslant m} \frac{1}{\mu_n} \psi_n^{(m)} \int_0^1 u^{(m)}(x)\,\psi_n^{(m)}(x)\,dx$$

converges and Parseval's theorem gives

$$\begin{aligned}|u^{(m)}|_0^2 &= \sum_{n\geq m}\left[\frac{1}{\sqrt{\mu_n}}\int_0^1 u^{(m)}(x)\,\psi_n^{(m)}(x)\,dx\right]^2\\ &= \sum_{n\geq m}\left[\frac{1}{\sqrt{\mu_n}}\mu_n\hat{u}_n\right]^2\\ &= \sum_{n\geq m}\mu_n\hat{u}_n^2 = \sum_{n\geq 0}\mu_n\hat{u}_n^2. \qquad (10)\end{aligned}$$

Also $f-u\in L^2[0,1]$, which entails that

$$|f-u|_0^2 = \sum_{n\geq 0}(\hat{u}_n-\hat{f}_n)^2. \qquad (11)$$

Substituting (10)–(11) into (2) our minimization problem becomes

$$\min_{\{\hat{u}_n\}}\left\{\lambda\sum_{n\geq 0}\mu_n\hat{u}_n^2+\sum_{n\geq 0}(\hat{u}_n-\hat{f}_n)^2\right\}. \qquad (12)$$

The solution is the function u with Fourier coefficients

$$\hat{u}_n = \frac{1}{1+\lambda\mu_n}\hat{f}_n \qquad (13)$$

and, hence,

$$|f-f_\lambda|_0^2 = \sum_{n\geq 0}\frac{\lambda^2\mu_n^2}{(1+\lambda\mu_n)^2}\hat{f}_n^2. \qquad (14)$$

We therefore see that the properties of $|f-f_\lambda|_0$ will depend strongly on the behavior of $\hat{f}_n$ and μ_n. Now we turn our attention to this problem.

3. The Behavior of the Fourier Coefficients

From the definition of $\hat{f}_n$ and ψ_n we get

$$\hat{f}_n = \int_0^1 \psi_n(x)\,f(x)\,dx = \frac{1}{\mu_n}\int_0^1 \psi_n^{(m)}(x)\,f^{(m)}(x)\,dx,$$

and integrating by parts m times we get for $f \in H^{2m}[0, 1]$

$$\hat{f}_n = \frac{1}{\mu_n} \sum_{j=0}^{m-1} (-1)^j \left[\psi_n^{(m-j-1)}(1) f^{(m+j)}(1) - \psi_n^{(m-j-1)}(0) f^{(m+j)}(0)\right]$$
$$+ \frac{(-1)^m}{\mu_n} \int_0^1 \psi_n(x) f^{(2m)}(x)\, dx. \tag{15}$$

The most convenient case for our purposes occurs when the first part vanishes. In this case we easily prove

THEOREM 2. *Let $f \in H^{2m}[0, 1]$ satisfy*

$$f^{(j)}(1) = f^{(j)}(0) = 0, \qquad j = m, ..., 2m-1. \tag{16}$$

Then

$$|f - f_\lambda|_0^2 \leqslant \lambda^2 |f^{(2m)}|_0^2. \tag{17}$$

Proof. According to (15) we have

$$\hat{f}_n = \frac{(-1)^m}{\mu_n} \int_0^1 \psi_n(x) f^{(2m)}(x)\, dx = \frac{(-1)^m}{\mu_n} \hat{f}_n^{(2m)}.$$

But $f^{(2m)} \in L^2[0, 1]$ and $\{\psi_n\}$ is an orthonormal basis for $L^2[0, 1]$. Thus Parseval's theorem gives

$$\sum_{n=m}^{\infty} [\hat{f}_n^{(2m)}]^2 \leqslant \sum_{n=0}^{\infty} [\hat{f}^{(2m)}]^2 = |f^{(2m)}|_0^2.$$

We now use this and (14) to get

$$\begin{aligned}
|f - f_\lambda|_0^2 &= \lambda^2 \sum_{n \geqslant 0} \frac{\mu_n^2 \hat{f}_n^2}{(1 + \lambda\mu_n)^2} \\
&= \lambda^2 \sum_{n \geqslant m} \frac{\mu_n^2 \hat{f}_n^2}{(1 + \lambda\mu_n)^2} \\
&= \lambda^2 \sum_{n \geqslant m} \frac{\mu_n^2 [\hat{f}_n^{(2m)}]^2}{(1 + \lambda\mu_n)^2\, \mu_n^2} \\
&\leqslant \lambda^2 \sum_{n \geqslant m} [\hat{f}_n^{(2m)}]^2 \\
&\leqslant \lambda^2 |f^{(2m)}|_0^2.
\end{aligned}$$

This concludes the proof. ∎

This is the result given by Locker and Prenter [4] since their condition becomes in this case $f \in R((D^m)^* (D^m))$ which entails the boundary conditions (16).

We now proceed to study the effect of the boundary on the convergence rates. From (15) we obtain

$$|\hat{f}_n| \leqslant \frac{1}{\mu_n} \sum_{j=0}^{m-1} |\psi_n^{(m-j-1)}(1)|\, |f_n^{(m+j)}(1)| + |\psi_n^{(m-j-1)}(0)|\, |f^{(m+j)}(0)| + \frac{1}{\mu_n} |\hat{f}_n^{(2m)}|. \tag{18}$$

We already know the behavior of the last term, we now turn our attention to the term involving the boundary conditions. To do this we must bound

$$\psi^{(k)}(0) \qquad \text{and} \qquad \psi^{(k)}(1), \qquad k = 0, ..., m-1.$$

Let us recall the following result due to Stone [11] on the behavior of the Green's function of a differential operator satisfying regular boundary conditions.

THEOREM 3 (Stone). *The residues of the Green's function for a regular differential system of order $n = 2m$ are a set of functions in x and y uniformly bounded for all x, y on* $(0, 1)$.

Let us apply this theorem to the following differential problem

$$D^{2m} V + (-1)^{m+1} \eta V = g, \tag{19}$$

$$V^{(j)}(0) = V^{(j)}(1) = 0, \qquad j = 0, ..., m-1. \tag{20}$$

Since it is known (cf. Naimark [7]) that the boundary conditions are regular, the Green's function has an expansion (cf. Birkhoff [2]) of the form

$$G(x, y; \eta) = (-1)^m \sum_{\nu \geqslant 0} \frac{\phi_\nu(x)\, \phi_\nu(y)}{\eta_\nu - \eta}, \tag{21}$$

where the ϕ_ν's are the eigenfunctions of $(-1)^m D^{2m}$ together with the boundary conditions (20) and the η_ν's their corresponding eigenvalues in increasing order. The ϕ_ν's are normalized by $|\phi_\nu|_0^2 = 1$.

We can now prove the following

LEMMA 4. *There exists a constant K depending only on m such that the eigenfunctions ψ_k, $k = m, m+1, \ldots$ of* (5) *normalized by $|\psi_k|_0^2 = 1$ satisfy*

$$\|\psi_k^{(m)}\|_\infty \leqslant K\sqrt{\mu_k}, \qquad k \geqslant m. \tag{22}$$

Proof. It is well known that (cf. [5])

$$\phi_\nu = \frac{1}{\sqrt{\mu_{\nu+m}}}\,\psi_{\nu+m}^{(m)}, \qquad \nu = 0, 1, 2, \ldots, \tag{23}$$

and

$$\eta_\nu = \mu_{\nu+m}.$$

Thus

$$G(x, y; \eta) = (-1)^m \sum_{\mu_{\nu+m}} \frac{1}{\mu_{\nu+m}} \frac{\psi_{\nu+m}^{(m)}(x)\,\psi_{\nu+m}^{(m)}(y)}{\mu_{\nu+m} - \eta} \tag{24}$$

and the residues of G at the poles $\mu_{\nu+m}$ are given by

$$R_\nu(x, y) = (-1)^{m+1} \frac{\psi_{\nu+m}^{(m)}(x)\,\psi_{\nu+m}^{(m)}(y)}{\mu_{\nu+m}}. \tag{25}$$

Applying now Theorem 3 we conclude the existence of $M > 0$ depending only on m such that

$$\left|\frac{\psi_k^{(m)}(x)\,\psi_k^{(m)}(y)}{\mu_k}\right| \leqslant M, \qquad x, y \in [0, 1]. \tag{26}$$

But (26) implies that

$$\left|\frac{\psi_k^{(m)}(x)}{\sqrt{\mu_k}}\right| \leqslant K, \qquad \forall x \in [0, 1],\ k = m, \ldots, \tag{27}$$

where $K = \sqrt{M}$ is independent of k. This concludes the proof. ∎

Let us now recall the following result for intermediate derivatives (cf. e.g., Schumaker [10, Theorem 2.4]).

THEOREM 5. *There exist constants C_j, $j = 1, \ldots, 2m-1$, depending only on m such that*

$$\|g^{(j)}\|_\infty \leqslant C_j(\varepsilon^{-j}\|g\|_\infty + \varepsilon^{2m-j}\|g^{(2m)}\|_\infty) \tag{28}$$

for any $g \in C^{2m}[0, 1]$ and any $0 < \varepsilon < \frac{1}{2}$.

We now use this result and Lemma 4 to prove the following result on $\|\psi_k^{(j)}\|_\infty$.

LEMMA 6. *There exist constants A_j, $j=0, ..., 2m-1$, depending only on m such that*

$$\|\psi_k^{(j)}\|_\infty \leqslant A_j \mu_k^{j/2m}, \qquad j=0, ..., m. \tag{29}$$

Proof. As we know, the eigenfunctions ψ_k belong to $C^\infty[0, 1]$. Thus

$$\psi_k^{(3m)}(x) = (-1)^m \mu_k \psi_k^{(m)}(x), \qquad \text{all} \quad x \in (0, 1).$$

Hence $\|\psi_k^{(3m)}\|_\infty = \mu_k \|\psi_k^{(m)}\|_\infty$ and from Lemma 4 we obtain

$$\|\psi_k^{(3m)}\|_\infty \leqslant K\mu_k^{3/2}. \tag{30}$$

We now apply Theorem 5 for $g = \psi_k^{(m)}$ and $\varepsilon = \mu_k^{-1/2m}$ to get

$$\|\psi_k^{(m+j)}\|_\infty \leqslant C_j(\mu_k^{j/2m} K\mu_k^{1/2} + \mu_k^{-(2m-j)/2m} K\mu_k^{3/2})$$

or

$$\|\psi_k^{(m+j)}\|_\infty \leqslant 2C_j K\mu_k^{(m+j)/2m}. \tag{31}$$

In particular, for $j=m$

$$\|\psi_k^{(2m)}\|_\infty \leqslant 2C_m K\mu_k. \tag{32}$$

We again apply Theorem 5 but this time to $g = \psi_k$ and $\varepsilon = \mu_k^{-1/2m}$. We obtain

$$\|\psi_k^{(j)}\|_\infty \leqslant C_j(\mu_k^{j/2m} \|\psi_k\|_\infty + \mu_k^{-(2m-j)/2m} \|\psi_k^{(2m)}\|_\infty). \tag{33}$$

Recalling that

$$\|\psi_k\|_\infty = \frac{1}{\mu_k} \|\psi_k^{(2m)}\|_\infty \leqslant 2KC_m,$$

(32) gives

$$\begin{aligned}\|\psi_k^{(j)}\|_\infty &\leqslant C_j(\mu_k^{j/2m} 2KC_m + \mu_k^{-(2m-j)/2m} 2KC_m \mu_k) \\ &\leqslant 4KC_m C_j \mu_k^{j/2m} \end{aligned} \tag{34}$$

which proves (29) for $A_0 = 2KC_m$, $A_{2m} = 2KC_m$, $A_j = 4KC_m C_j$, $j=1, ...,$ $2m-1$. ∎

We are now in a position to give the exact behavior of the $\hat{f}_i$'s.

THEOREM 7. *Let $f \in H^{2m}[0, 1]$. Then there exist constants $A_0, \ldots, A_{2m}$ depending only on m such that*

$$|\hat{f}_i| \leqslant \sum_{j=0}^{m-1} [|f^{(m+j)}(0)| + |f^{(m+j)}(1)|] A_{m-j-1} \mu_i^{-(m+j+1)/2m} + o(\mu_i^{-1}). \tag{35}$$

Moreover, for f satisfying the boundary conditions

$$f^{(j)}(0) = f^{(j)}(1) = 0, \qquad j = m, \ldots, k,\ m \leqslant k < 2m - 1,$$

then

$$|\hat{f}_i| \leqslant \mu_i^{-(k+2)/2m} \sum_{j=1}^{2m-1-k} [|f^{(k+j)}(1)| + |f^{(k+j)}(0)|] \times A_{2m-j-1} \mu_i^{-(j-1)/2m} + o(\mu_i^{-1}) \tag{36}$$

($k = m - 1$ means that either $f^{(m)}(0) \neq 0$ or $f^{(m)}(1) \neq 0$).

Proof. From (15) we have

$$\hat{f}_i = \frac{1}{\mu_i} \sum_{j=0}^{m-1} (-1)^j [\psi_i^{(m-j-1)}(1) f^{(m+j)}(1) - \psi_i^{(m-j-1)}(0) f^{(m+j)}(0)] + \frac{(-1)^m}{\mu_i} \hat{f}_i^{(2m)}.$$

Taking absolute values and using Lemma 6 we get

$$\begin{aligned} |\hat{f}_i| &\leqslant \frac{1}{\mu_i} \sum_{j=0}^{m-1} [|f^{(m+j)}(1)| + |f^{(m+j)}(0)|] \|\psi_i\|_\infty^{m-j-1} + \frac{|\hat{f}_i^{(2m)}|}{\mu_i} \\ &\leqslant \frac{1}{\mu_i} \sum_{j=0}^{m-1} [|f^{(m+j)}(1)| + |f^{(m+j)}(0)|] A_{m-j-1} \mu_i^{(m-j-1)/2m} + \frac{|\hat{f}_i^{(2m)}|}{\mu_i} \\ &\leqslant \sum_{j=0}^{m-1} [|f^{(m+j)}(1)| + |f^{(m+j)}(0)|] A_{m-j-1} \mu_i^{-(m+j+1)/2m} + \frac{1}{\mu_i} |\hat{f}_i^{(2m)}|. \end{aligned} \tag{37}$$

This proves (34) since $f \in H^{2m}[0, 1]$ implies

$$\sum_{i \geqslant 0} |\hat{f}_i^{(2m)}|^2 < +\infty,$$

hence $|\hat{f}_i^{(2m)}| \to 0$ as $i \to \infty$. Equation (36) is obtained from (35) using the boundary conditions. ∎

As we can observe from (36), the rate of decay of the $\hat{f}_i$'s to zero is strongly affected by the values of $f^{(j)}$ at the boundary. It is clear for instance that the smoothness of f is less important than the values of $f^{(j)}$ at the end points of the interval. Thus, for example, even for a very smooth function ($f \in H^{2m}$), the rate of decay will be $\mu_i^{-(m+1)/2m}$ if $f^{(m)}(0) \neq 0$ or $f^{(m)}(1) \neq 0$. This fact is going to be of crucial importance in the study of convergence rates for Tikhonov regularization as we shall see in the next section.

4. Convergence Rates for the Tikhonov Regularizer

Let us now recall the expansion for the error that we have obtained in Section 2, namely,

$$|f - f_\lambda|_0^2 = \sum_{n \geqslant m} \frac{\lambda^2 \mu_n^2}{(1 + \lambda \mu_n)^2} \hat{f}_n^2.$$

Suppose that $f \in H^{2m}[0, 1]$ and for some $m \leqslant k < 2m - 1$ we have

$$f^{(j)}(0) = f^{(j)}(1) = 0, \qquad j = m, ..., k.$$

Then by Theorem 7 there exists a constant $B > 0$ such that

$$\hat{f}_n^2 \leqslant B \mu_n^{-(k+2)/m}$$

$$|f - f_\lambda|_0^2 \leqslant \sum_{n \geqslant m} \frac{\lambda^2 \mu_n^2}{(1 + \lambda \mu_n)} (B \mu_n^{-(k+2)/m})$$

$$\leqslant B \lambda^{(k+2)/m} \sum_{n \geqslant m} \frac{(\lambda \mu_n)^{(2m-k-2)/m}}{(1 + \lambda \mu_n)^2}. \tag{38}$$

Now we use Birkhoff results on the behavior of the μ_n's (cf. [2]). That is, there exist constants $\alpha, \beta > 0$ (independent of n) such that

$$\alpha n^{2m} \leqslant \mu_n \leqslant \beta n^{2m}. \tag{39}$$

This gives

$$|f - f_\lambda|_0^2 \leqslant B \lambda^{(k+2)/m} \sum_{n \geqslant m} \frac{(\lambda \beta n^{2m})^{(2m-k-2)/m}}{(1 + \alpha \lambda n^{2m})^2}$$

$$\leqslant B \left(\frac{\beta}{\alpha}\right)^{(2m-k-2)/m} \lambda^{(k+2)/m} \sum_{n \geqslant m} \frac{(n\theta)^{4m-2k-4}}{(1 + (n\theta)^{2m})^2},$$

with $\theta = (\alpha\lambda)^{1/2m}$. But $w(x) = x^{4m-2k-4}/(1+x^{2m})^2$ is increasing for $0 \leqslant x < x_0$ and decreasing for $x_0 < x$ where

$$x_0 = \sqrt{\frac{4m-2k-4}{2k+4}}. \tag{40}$$

Hence

$$\theta \sum_{n \geqslant m} w(n\theta) = \theta \left[\sum_{n=m}^{p} w(n\theta) + \sum_{n=p+1}^{\infty} w(n\theta) \right],$$

where p is such that $p\theta \leqslant x_0 < (p+1)\,\theta$. Then

$$\theta \sum_{n=p+1}^{\infty} w(n\theta) \leqslant \int_{p\theta}^{\infty} w(x)\, dx$$

and

$$\theta \sum_{n=m} w(n\theta) \leqslant \int_{0}^{p\theta} w(x)\, dx.$$

Moreover,

$$w(p\theta) \leqslant w(x_0) = \text{constant}.$$

Thus, finally,

$$\begin{aligned} \sum_{n \geqslant m} w(n\theta) &\leqslant [I_{m,k} + \theta w(x_0)] \frac{1}{\theta} \\ &\leqslant [I_{m,k} + \lambda^{1/2m}\alpha^{1/2m} w(x_0)]\, \alpha^{-1/2m}\lambda^{-1/2m}, \end{aligned} \tag{41}$$

where

$$I_{m,k} = \int_0^\infty \frac{x^{4m-2k-4}}{(1+x^{2m})^2}\, dx \tag{42}$$

is bounded since $4m > (4m-2k-4)+2$. We thus get

$$|f - f_\lambda|_0^2 \leqslant B \left(\frac{\beta}{\alpha}\right)^{2m-k-2} \alpha^{-1/2m} [I_{m,k} + w(x_0)\, \lambda^{1/2m}]\, \lambda^{(2k+3)/2m}. \tag{43}$$

Let us now examine the error in the mth derivative. We have

$$|f - f_\lambda|_m^2 = \sum_{n \geqslant m} \frac{\lambda^2 \mu_n^2}{(1+\lambda\mu_n)^2} \hat{f}_n^2 \mu_n.$$

Using (37) again, we get

$$|f-f_\lambda|_m^2 \leqq B \sum_{n\geqslant m} \frac{\lambda^2\mu_n^2}{(1+\lambda\mu_n)^2}\,\mu_n^{(m-k-2)/m}$$
$$\leqslant B\lambda^{(k+2-m)/m} \sum_{n\geqslant m} \frac{(\lambda\mu_n)^{(3m-k-2)/m}}{(1+\lambda\mu_n)^2}. \tag{44}$$

Now replacing Birkhoff inequalities for the eigenvalues gives

$$|f-f_\lambda|_m^2 \leqslant B\lambda^{(k+2-m)/m} \sum_{n\geqslant m} \frac{(\lambda\beta n^{2m})^{(3m-k-2)/m}}{(1+\lambda\alpha n^{2m})^2}$$
$$\leqslant B\left(\frac{\beta}{\alpha}\right)^{(3m-k-2)/m} \lambda^{(k+2-m)/m} \sum_{n\geqslant m} \frac{(\theta n)^{6m-2k-4}}{(1+(\theta n)^{2m})^2}, \tag{45}$$

where $\theta=(\alpha\lambda)^{1/2m}$. This gives, after some algebra,

$$|f-f_\lambda|_m^2 \leqslant B\left(\frac{\beta}{\alpha}\right)^{(3m-k-2)/m} \alpha^{-1/2m}[J_{m,k}+\tilde{\omega}(\tilde{X}_0)^{1/2m}]\,\lambda^{(2(k-m)+3)/2m}. \tag{46}$$

Here

$$\tilde{\omega}(X)=\frac{X^{6m-2k-4}}{(1+X^{2m})^2},$$
$$\tilde{\omega}'(X_0)=0, \tag{47}$$
$$J_{m,k}=\int_0^\infty \frac{X^{6m-2k-4}}{(1+X^{2m})^2}\,dX$$

is convergent since $4m>6m-2k-4+2$ $(m\leqslant k)$.

Let us now state the main result of this paper.

THEOREM 8. *Let* $f\in H^{2m}[0,1]$ *be such that*

$$f^{(j)}(0)=f^{(j)}(1)=0, \qquad j=m,\ldots,k,$$

where $m-1<k\leqslant 2m-2$ *is a given integer.* (*If* $k=m-1$ *none of these conditions are satisfied.*) *Then there exist constants* $D_{j,k}$, $0\leqslant j\leqslant m$, *independent of* λ *such that for* $\lambda<(1/2)^{2m}$ *we have*

$$|f-f_\lambda|_j^2 \leqslant D_{j,k}\lambda^{(2k+3-2j)/2m}, \qquad j=0,1,\ldots,m.$$

Proof. From (43) and (46) there exist constants $R_{0,k}$, $R_{m,k}$ depending only on m, k such that

$$|f-f_\lambda|_0^2 \leqslant R_{0,k}\lambda^{(2k+3)/2m} \tag{48}$$

$$|f-f_\lambda|_m^2 \leqslant R_{m,k}\lambda^{(2(k-m)+3)/2m} \tag{49}$$

for $\lambda < (1/2)^{2m} < 1$ where

$$R_{0,k} = B\left(\frac{\beta}{\alpha}\right)^{2m-k-2} \alpha^{-1/2m}[I_{m,k} + w(x_0)]$$

$$R_{m,k} = B\left(\frac{\beta}{\alpha}\right)^{(3m-k-2)/m} \lambda^{-1/2m}[J_{m,k} + \tilde{w}(\tilde{x}_0)].$$

We now use Agmon's theorem (cf. Theorem 5) to conclude the existence of $A_0, ..., A_m$ depending only on m such that

$$|f-f_\lambda|_j^2 \leqslant A_j(\varepsilon^{-j}|f-f_\lambda|_0^2 + \varepsilon^{m-j}|f-f_\lambda|_m^2)$$

for $0 < \varepsilon < \frac{1}{2}$. Let us take $\varepsilon = \lambda^{1/m}$ and get

$$\begin{aligned}|f-f_\lambda|_j^2 &\leqslant A_j(\lambda^{-j/m}R_{0,k}\lambda^{(2k+3)/2m} + \lambda^{(m-j)/m}R_{m,k}\lambda^{(2(k-m)+3)/2m})\\ &\leqslant A(R_{0,k}+R_{m,k})\,\lambda^{(2(k-j)+3)/2m}\end{aligned}$$

which proves the theorem for $D_{j,k} = A_j(R_{0,k} + R_{m,k})$. ∎

The effect of the values of f at the boundary is clearly established in this theorem; for instance, if $f \in H^{2m}[0, 1]$ does not satisfy any special condition, the error in the function is $|f-f_\lambda|_0^2 \leqslant \text{constant } \lambda^{1+1/2m}$. However, if $f^{(m)}(0) = f^{(m)}(1) = 0$, $|f-f_\lambda|_0^2 \leqslant \text{constant } \lambda^{1+3/2m}$ without any additional hypothesis on the smoothness of f.

In smoothing by spline functions it is of interest to study the error in the smoothing process. One of the error terms is the Integrated Mean Square Error (IMSE). We can connect this IMSE to $|f-f_\lambda|_0^2$ by observing the f_λ is "an approximation" to $S_{n,\lambda}$, the smoothing spline defined as (cf. [13, 14]) the solution to

$$\min_{u \in H^m[0,1]} \left\{\lambda |u|_m^2 + \frac{1}{n}\sum_{i=1}^{n} [u(t_i) - f(t_i)]^2\right\},$$

where

$$t = \frac{(2i-1)}{2}\frac{1}{n}, \qquad i = 1, ..., n.$$

It is shown in [8] that the IMSE in the smoothing of noisy data is given by

$$\text{IMSE} = |f - S_{n,\lambda}|_0^2 \approx \frac{1}{n} \sum_{i=1}^{n} [f(t_i) - S_{n,\lambda}(t_i)]^2.$$

If we consider $|f - f_\lambda|_0^2$ as a good approximation to $|f - S_{n,\lambda}|_0^2$ (cf. [13]) we conclude that

$$\text{IMSE} \approx RD_{0,k}\lambda^{(2k+3)/2m}.$$

This extends the results of Rice and Rosemblatt (cf. [9]) for cubic smoothing splines ($m = 2$) and allows us to expect for general m the result

$$|f - S_{n,\lambda}|_j^2 = O(\lambda^{(2(k-j)+3)/2m}).$$

Acknowledgments

This work was partially developed while visiting the Department of Mathematics and Computer Science of the University of Miami. The author thanks the Chairman of the Department, Dr. Shair Ahmad; the colleagues; and all the staff of the Department for the hospitality he received during his entire stay.

The author also thanks two unknown referees for helping to improve the manuscript by suggesting many changes.

References

1. S. Agmon, "Lectures on Elliptic Boundary Value Problems," Van Nostrand Mathematical Studies, Princeton, NJ, 1965.
2. G. D. Birkhoff, Bundary value and expansion problems of ordinary linear differential equations, *Trans. Amer. Math. Soc.* **9** (1908), 373–395.
3. P. Craven and G. Wahba, Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by cross-validation, *Numer. Math.* **31** (1979), 377–403.
4. J. Locker and P. M. Prenter, Regularization with differential operators. I. General theory, *J. Math. Anal. Appl.* **74** (1980), 504–529.
5. G. G. Lorentz, "Approximation of Functions," Holt, Rinehart & Winston, New York, 1966.
6. A. A. Melkman and C. A. Micchelli, Spline spaces are optimal for L^2 n-width, *Illinois J. Math.* **22** (4) (1978).
7. M. A. Naimark, "Linear Differential Operators," Vol. I, Ungar, New York, 1967.
8. D. L. Ragozin, Error bounds for derivative estimates based on spline smoothing of exact or noisy data, *J. Approx. Theory* **37** (1983), 335–355.
9. J. Rice and M. Rosenblatt, Smoothing splines: Regression, derivatives and deconvolution, *Ann. Statist.* **11** (1983), 141–156.

10. L. L. SCHUMAKER, "Spline Functions: Basic Theory," Wiley, New York, 1981.
11. M. L. STONE, A comparison of the series of Fourier and Birkhoff, *Trans. Amer. Math. Soc.* **28** (1926), 695–761.
12. A. N. TIKHONOV AND V. Y. ARSENIN, "Solutions of Ill-Posed Problems," Wiley, New York, 1977.
13. F. I. UTRERAS, Optimal smoothing of noisy data using spline functions, *SIAM J. Sci. Statist. Comput.* **1** (1981).
14. G. WAHBA, Smoothing and ill-posed problems, *in* "Solution Methods for Integral Equations with Applications" (Michel Goldberg, Ed.), Plenum, New York, 1979.