JOURNAL OF COMPUTATIONAL PHYSICS 42, 141–151 (1981)

Computational Experience with the Spectral Smoothing Method for Differentiating Noisy Data

M. L. BAART

National Research Institute for Mathematical Sciences of the CSIR, P.O. Box 395, Pretoria, South Africa

Received August 20, 1980; revised December 8, 1980

When applied to non-exact (noisy) data, numerical methods for calculating derivatives, in particular derivatives of order higher than the first, based on model functions fitted to exact data, become unsatisfactory . The spectral smoothing method of Anderssen and Bloomfield, developed to solve this problem, entails calculation of a smoothing parameter and the choice of an optimal-order Sobolev norm that is used as regularizer. This method is used to differentiate, smooth and integrate noisy data. A likelihood function is minimized to determine the smoothing parameter. We present numerical results suggesting that this function can be jointly minimized with respect to the smoothing parameter and the order of the regularizing norm, thus yielding a fully automatic numerical differentiation procedure.

1. INTRODUCTION

The problem of finding the nth derivative $f(\cdot)$ of a function $g(\cdot)$, i.e.,

$$f(t) = g^{(n)}(t), \qquad t_0 \leqslant t \leqslant T, \quad n \geqslant 1, \tag{1.1}$$

can be formulated in terms of integrals. Assuming that we know the lower-order derivatives at $t = t_0$, we may use Taylor's theorem with remainder to define

$$G(t) := g(t) - \sum_{k=0}^{n-1}{}' (t - t_0)^k \, g^{(k)}(t_0)/k!$$

$$= \int_{t_0}^{t} (t - \tau)^{n-1} \, g^{(n)}(\tau)/(n-1)! \, d\tau,$$

with

$$G^{(n)}(t) = g^{(n)}(t) \qquad \text{and} \qquad G^{(k)}(t_0) = 0, \; 0 \leqslant k \leqslant n-1. \tag{1.2}$$

The problem of solving (1.1) is therefore transformed to that of solving the Volterra integral equation of the first kind

$$D_n f \,|\, (t) := \int_{t_0}^{t} (t - \tau)^{n-1} \, f(\tau)/(n-1)! \, d\tau = G(t), \tag{1.3}$$

0021-9991/81/070141-11S02.00/0

since (1.2) implies that $G(\cdot)$ and $g(\cdot)$ are equivalent modulo a polynomial of degree $(n-1)$. This is a special form of the more general Fredholm integral equation of the first kind

$$Kf\,|\,(t) := \int_{t_0}^{T} K(t,\tau)\, f(\tau)\, d\tau = G(t), \qquad t_0 \leqslant t \leqslant T, \tag{1.4}$$

with the difference kernel

$$\begin{aligned} K(t,\tau) &= (t-\tau)^{n-1}/(n-1)!, && t_0 \leqslant \tau \leqslant t \leqslant T, \\ &= 0, && t_0 \leqslant t < \tau \leqslant T. \end{aligned} \tag{1.5}$$

The problem of solving (1.3) or (1.4) for the unknown function $f(\cdot)$ is known to be ill-posed (cf. [3, 7]) in the sense that, even if a solution exists, it need neither be unique nor depend continuously on the right-hand side (data). This implies that small inaccuracies (noise) in the data can lead to calculated "solutions" that do not approximate the exact solution in any sense.

The degree of smoothness of the kernel plays an important role in numerical procedures for solving integral equations of the first kind (cf. [4, 6]). For example, if the kernel approximates the Dirac delta generalized function, discretization of (1.4) yields a diagonally dominant coefficient matrix, and inversion presents no problems. If the kernel is constant, however, the discretization matrix will have a rank of one since all its rows are identical. The value of n (the order of the derivative) determines the smoothness of the kernel (1.5): the problem will become progressively more ill-conditioned as n increases.

2. Regularization

If an ill-posed problem such as (1.3) is to be regularized, this means simply that it is to be replaced by a problem that is well-posed (has a unique solution that depends continuously on the data) and is such that its solution $\tilde{f}(\cdot)$ approximates the function $f(\cdot)$ in some satisfactory way, usually in L^2-norm, i.e.,

$$\|\tilde{f} - f\|^2 := \int_{t_0}^{T} (\tilde{f}(\tau) - f(\tau))^2\, d\tau < \varepsilon, \tag{2.1}$$

for given small $\varepsilon > 0$. We assume that $f(\cdot)$ is an element of the Sobolev space $H^m[t_0, T]$, i.e.,

$$\|f\|_m^2 := \sum_{k=0}^{m} \int_{t_0}^{T} (f^{(k)}(\tau))^2\, d\tau < \infty, \tag{2.2}$$

for some positive integer value of m.

One way of regularizing (1.3) is discussed in [7], where it is shown that for values of α in a restricted interval (β_1, β_2) dependent on ε, the (unique) solution $f_\alpha(\cdot)$ of the optimization problem

$$\inf_{f \in H^m} \{\|D_n f - G\|^2 + \alpha \|f\|_m^2\}, \qquad 0 < \beta_1(\varepsilon) < \alpha < \beta_2(\varepsilon) < 1, \tag{2.3}$$

will have the desired property (2.1). Standard results of the calculus of variations then require $f_\alpha(\cdot)$ to be a solution of the integro-differential equation

$$D_n^*(D_n f - G) + \alpha \sum_{k=0}^{m} (-1)^k f^{(2k)} = 0, \tag{2.4}$$

where D_n^* denotes the adjoint operator of D_n, and in this case has the form

$$D_n^* g \,|\, (t) := \int_t^T (\tau - t)^{n-1} g(\tau)/(n-1)!\, d\tau. \tag{2.5}$$

The boundary conditions associated with (2.4) are

$$f^{(m+i)}(t_0) + (-1)^i f^{(m-i)}(t_0) = f^{(m+i)}(T) + (-1)^i f^{(m-i)}(T) = 0, \qquad 0 \leqslant i \leqslant m-1. \tag{2.6}$$

Eq. (2.4) is no longer ill-posed. Cullum [3] shows, for instance, that with $n = 1$, the equation can, using Green's functions, be transformed to a (well-posed) integral equation of the second kind. However, the process of solving this transformed equation for a given value of α, although numerically stable, requires more computations than the procedure described below (cf. [1]), and the optimal value of α still has to be estimated in some way.

3. The Anderssen–Bloomfield (A–B) Approach

Equation (2.4) can be expressed in the form

$$\int_t^T \int_{t_0}^{\tau_2} (\tau_2 - t)^{n-1} (\tau_2 - \tau_1)^{n-1} f(\tau_1)/((n-1)!)^2\, d\tau_1\, d\tau_2 + \alpha \sum_{k=0}^{m} (-1)^k f^{(2k)}(t)$$
$$= \int_t^T (\tau - t)^{n-1} G(\tau)/(n-1)!\, d\tau. \tag{3.1}$$

Differentiation of (3.1) $2n$ times with respect to the variable t yields the $(2m + 2n)$th-order differential equation

$$f(t) + \alpha \sum_{k=0}^{m} (-1)^{n+k} f^{(2n+2k)}(t) = g^{(n)}(t), \tag{3.2}$$

with the boundary conditions (additional to (1.2) and (2.6))

$$\sum_{k=0}^{m} (-1)^k f^{(2k+i)}(T) = \sum_{k=0}^{m} (-1)^{n-k} f^{(n+2k+i)}(t_0) - g^{(i)}(t_0)$$
$$= 0, \qquad 0 \leqslant i \leqslant n-1. \tag{3.3}$$

Additional terms in (3.1), used in [1–3] to enforce boundary conditions, have been omitted for convenience, as they do not affect (3.2) and (3.3).

To simplify the series representation of the solution and data, the function $g(\cdot)$ underlying the data is detrended, i.e., we assume that the data at the endpoints are accurate, and subtract a linear function from $g(\cdot)$ to ensure that $g(t_0) = g(T) = 0$. Where necessary, corresponding adjustments will have to be made to solutions. It follows that $g(\cdot)$ may be expanded in a Fourier sine series

$$g(t) = \sum_{j=1}^{\infty} \gamma_j \sin \phi_j t, \tag{3.4}$$

where

$$\phi_j := \pi j/(T - t_0).$$

By substitution in (3.2) it is possible to show formally that the solution is given by

$$f(t) = \sum_{j=1}^{\infty} \gamma_j w_j \phi_j^n \sin\left(\phi_j t + \frac{n\pi}{2}\right), \tag{3.5}$$

where

$$w_j^{-1} := 1 + \alpha \phi_j^{2n} \sum_{k=0}^{m} \phi_j^{2k}. \tag{3.6}$$

The regularizing procedure therefore produces weighting factors w_j which act on the Fourier coefficients γ_j of the data. The convergence of the expansion (3.4), and therefore the number of terms to be retained for calculating approximate solutions, is adversely affected by the presence of noise in the data. The weight-function serves to counteract this effect.

When $g(\cdot)$ is represented by N discrete equispaced point values, we can construct a finite Fourier series corresponding to (3.4) by assuming an odd extension of the data, with the regularized nth derivative given by the corresponding finite form of (3.5). Fast Fourier transforms (FFTs) can then be used to calculate γ_j, which makes the algorithm computationally efficient. The value of the regularizing constant α to be used in the computations may be selected, as is shown in [1, 2], by interpreting it as a parameter in a statistical model, and calculating its maximum likelihood estimate. This is found by minimizing the following likelihood function with respect to α:

$$\Phi(\alpha; m, n) := (N-1) \ln\left(\sum_{j=1}^{N-1} \tilde{\gamma}_j^2 (1 - \tilde{w}_j)\right) - \sum_{j=1}^{N-1} \ln(1 - \tilde{w}_j). \tag{3.7}$$

Here the $\tilde{\gamma}_j$ are the FFT-calculated approximations to the Fourier coefficients γ_j of the data, and the $\tilde{w}_j$ are the finite equivalents of w_j, i.e., with ϕ_j replaced by

$$\tilde{\phi}_j := \pi j/(N-1).$$

It is in this context that the value of α used in the calculations is considered to be optimal. The likelihood function (3.7) is derived under the assumption of uncorrelated errors each following the same Gaussian distribution. However, as is indicated in [8], the expression (3.7) is asymptotically valid for large N under considerably relaxed conditions. In particular, the errors need only have zero mean, constant finite variance and be uncorrelated. Constants in (3.7) that do not influence the optimization have been omitted, so that this expression differs slightly from those used in [1, 2, 5].

Although Eqs. (3.2) to (3.7) inclusive were derived for $n \geqslant 1$, the same approach is feasible for $n = 0$ (as noted in [1]), in which case $D_0 \equiv I$, the identity operator, and an optimally smoothed form of the data is obtained. For negative values of n, say $n = -\nu$, we can interpret $D_{-\nu} f$ as $f^{(\nu)}(\cdot)$ in (2.3), and $f(\cdot)$ as the νth indefinite integral of the data $g(\cdot)$.

Although integration is a smoothing process, and noise in the data will seldom cause problems, it is of some interest to note that the A–B procedure can be used to differentiate, smooth and integrate noisy data optimally (in the present context). For example, the acceleration, smoothed velocity and displacement of some motion represented by velocity measurements may be calculated by using $n = 1, 0$ and -1, respectively.

Where integration is performed, some care has to be taken with adjustments of solutions necessitated by detrending the data, and in the choice of m, the order of the regularizing Sobolev norm. These points will be discussed in more detail in Section 5. Also, although the boundary conditions (1.3), (2.6) and (3.3) have to be satisfied theoretically, this will seldom (if ever!) be true in practice. Function estimates near the endpoints of the data interval should therefore be treated with caution.

4. Applications

In the following applications the data represent the functions

$$g_{ij}(t) := g_i(t) + \varepsilon_j r(t), \qquad 0 \leqslant t \leqslant 1, \quad j = 0, 1, 2, 3, \tag{4.1}$$

where $g_i(t)$ is a given test function, $r(t)$ is an error function uniformly distributed over $(-1, 1)$, and

$$\varepsilon_j / g_{i\max} := \begin{cases} 0.0, & j = 0, \\ 0.01, & j = 1, \\ 0.05, & j = 2, \\ 0.1, & j = 3, \end{cases}$$

where

$$g_{i\max} := \max_{0 \le t \le 1} |g_i(t)|.$$

This allows for maximum errors in the data of 0, 1, 5, and 10%, respectively. The two test functions considered are

$$g_1(t) := \sin \pi t,$$
$$g_2(t) := 3et^3 \ln(1/t).$$

The constant $3e$ in $g_2(\cdot)$ is introduced to ensure that $g_{i\max} = 1$. Detrending is unnecessary since $g_i(0) = g_i(1) = 0$.

For the function $g_1(\cdot)$ derivatives of all orders are bounded, while for $g_2(\cdot)$ only the first and second derivatives are bounded.

As in the preceding sections, n determines the operation to be performed on the data, m is the order of the regularizing Sobolev norm, and N is the number of data points.

The values of n, m and N belonged to the sets $\{-2, -1, 0, 1, 2\}$, $\{1, 2, 3, 4, 5, 6\}$ and $\{26, 51, 101, 201\}$, respectively. To exclude endpoint effects, the first and last truncated tenth of the N data points used in solution calculations were ignored in error calculations.

The following two criteria for the accuracy of the computed solutions were used. Let S denote the set of $\tilde{N}$ data points $\{t_i\}$ used for error calculations, $f(\cdot)$ the exact solution for no noise and with maximum absolute value $f_{\max}$, and $\tilde{f}(\cdot)$ the computed solution. Then

$$E_2 := \left(\frac{1}{\tilde{N}} \sum_{t_i \in S}' (f(t_i) - \tilde{f}(t_i))^2\right)^{1/2} \Big/ f_{\max},$$
$$E_\infty := \max_{t_i \in S} |f(t_i) - \tilde{f}(t_i)| / f_{\max}.$$

All numerical calculations were done on a CDC Cyber 174 computer. The notation aEb denotes $a \cdot 10^b$.

5. Discussion of Results

As it is impossible to give a full account of the results obtained, only general trends observed will be discussed, and the effect emphasized of changes in the order of the regularizing norm, which in practice will be the only parameter choice left to the user of the algorithm. This implies that we are in effect looking for an optimal Sobolev regularizer for the numerical differentiation problem. A study by Cullum [4] of the effects of using various smoothing norms for ill-posed problems shows the importance of estimating how appropriate any proposed regularizer actually is.

TABLE I

Effect of Changes in N on Errors E_2 and E_∞

(g_{ij}, m, n)	N	E_2	E_∞
$(g_{12}, 1, 0)$	26	0.308 E − 1	0.493 E − 1
	51	0.283 E − 1	0.493 E − 1
	101	0.222 E − 1	0.417 E − 1
	201	0.159 E − 1	0.404 E − 1
$(g_{21}, 3, 1)$	26	0.111 E − 1	0.387 E − 1
	51	0.456 E − 2	0.151 E − 1
	101	0.513 E − 2	0.115 E − 1
	201	0.320 E − 2	0.808 E − 2
$(g_{23}, 2, 3)$	26	0.128	0.428
	51	0.713 E − 1	0.234
	101	0.574 E − 1	0.188
	201	0.479 E − 1	0.128

5.1. For most functions g_{ij} with fixed smoothing norm and for a specified operation (i.e., values of n), both error measures E_2 and E_∞ decreased with an increase in the number of data points. Some results are given in Table I, and the general nature of these could have been expected from the nature of the estimation of $\tilde{\Phi}$ and $\tilde{\alpha}$, the optimal values of the likelihood function and the smoothing parameter, respectively. Note that an increase in the number of observations may imply that E_2 and/or E_∞ for $f \equiv g_i$ and $\tilde{f} \equiv g_{ij}$ increase, so that an improvement in the calculated solution need not always follow, even though the error amplitude ε_j is kept constant. However, in contrast to the instability of finite difference or curve-fitting methods when intervals between noisy data decrease, error fluctuations are random in nature, and stable.

5.2. The optimal values $\tilde{\alpha}$ and $\tilde{\Phi}$ of the smoothing parameter and the likelihood function, for each function g_{ij} represented by a fixed number of data points, depended strongly on $(m + n)$, i.e., the norm of the highest derivative in (2.2) dominated the Sobolev norm for these examples, as shown in Table II. For fixed $(m + n)$, only slight

TABLE II

Dependence of $\tilde{\alpha}$ and $\tilde{\Phi}$ on $(m + n)$

(g_{ij}, N)	$m + n$	$\tilde{\alpha}$	$\tilde{\Phi}$
$(g_{12}, 51)$	4	0.714 E − 8	−0.169 E3
$(g_{13}, 201)$	6	0.524 E − 10	−0.151 E3
$(g_{21}, 26)$	3	0.733 E − 9	−0.125 E3
$(g_{22}, 101)$	2	0.316 E − 6	−0.256 E3
$(g_{23}, 51)$	5	0.537 E − 11	−0.886 E2

variations in $\tilde{\alpha}$ occurred, e.g., in the case g_{22}, $n = 101$, $m + n = 2$, the values of $\tilde{\alpha}$ for $n = -1$, $m = 3$ and $n = -2$, $m = 4$ differed by less than 1%. This seems to indicate that (2.2) can be replaced by $\|f^{(m)}\|^2$ as regularizer. The weights $\tilde{w}_j$ will of course be different. The effect on the accuracy of the solutions has not been investigated. Although the applicability of the A–B algorithm to numerical integration is academic rather than practical, this phenomenon means that, for $n = -v$, m has to be so chosen that $(m + n) = (m - v) > 0$; otherwise the optimization of (3.7) leads to problems. For instance, with $m + n = 0$ and using (3.6) we have the "equivalent" form of (3.7)

$$\Phi(\alpha; 0, 0) = (N-1) \ln \left(\sum_{j=1}^{N-1} \alpha \tilde{\gamma}_j^2/(1+\alpha) \right) - \sum_{j=1}^{N-1} \ln(\alpha/(1+\alpha))$$

$$= (N-1) \ln \sum_{j=1}^{N-1} \tilde{\gamma}_j^2,$$

which is independent of α, and therefore cannot be optimized. This appeared in the numerical experiments as spurious convergence to optimal values of $\tilde{\alpha} \geqslant 1$, and large positive values of $\tilde{\Phi}$.

5.3. When the noise amplitude was the only parameter allowed to vary, the optimal value $\tilde{\alpha}$ increased with increasing noise amplitude ε_j, which means that, as is to be expected, more smoothing is required for larger inaccuracies in the data (Table III).

On the other hand, when n was allowed to vary, for a specific data set (g_{ij}, N) with fixed regularizer, the optimal value $\tilde{\alpha}$ of the smoothing parameter decreased with increase in n. This is misleading, because although $\tilde{\alpha}$ might decrease, the increase in n dominates the weights $\tilde{w}_j$, so that more smoothing of the Fourier coefficients of the data does in fact take place for the optimal calculation of higher derivatives. Figure 1 shows the extent of the spectral smoothing for the function g_{12}, with the third-order Sobolev norm as regularizer and 51 representative data points. When one integration of the data is performed ($n = -1$), very little smoothing is required compared with that needed to compute the second derivative ($n = 2$), even though the optimal smoothing parameter $\tilde{\alpha}$ is now much smaller.

TABLE III

Changes in $\tilde{\alpha}$ with Increased Noise

(g_{ij}, m, n, N)	$\tilde{\alpha}$ $(j=0)$	$\tilde{\alpha}$ $(j=1)$	$\tilde{\alpha}$ $(j=2)$	$\tilde{\alpha}$ $(j=3)$
$(g_{1j}, 2, 1, 51)$	0.199 E − 36	0.652 E − 8	0.964 E − 7	0.296 E − 6
$(g_{1j}, 1, 2, 101)$	0.300 E − 38	0.360 E − 8	0.512 E − 7	0.135 E − 6
$(g_{2j}, 3, 2, 51)$	0.160 E − 19	0.264 E − 14	0.421 E − 12	0.537 E − 11
$(g_{2j}, 2, 1, 101)$	0.299 E − 39	0.920 E − 10	0.308 E − 8	0.149 E − 7
$(g_{2j}, 3, 2, 201)$	0.207 E − 25	0.151 E − 15	0.579 E − 13	0.518 E − 12

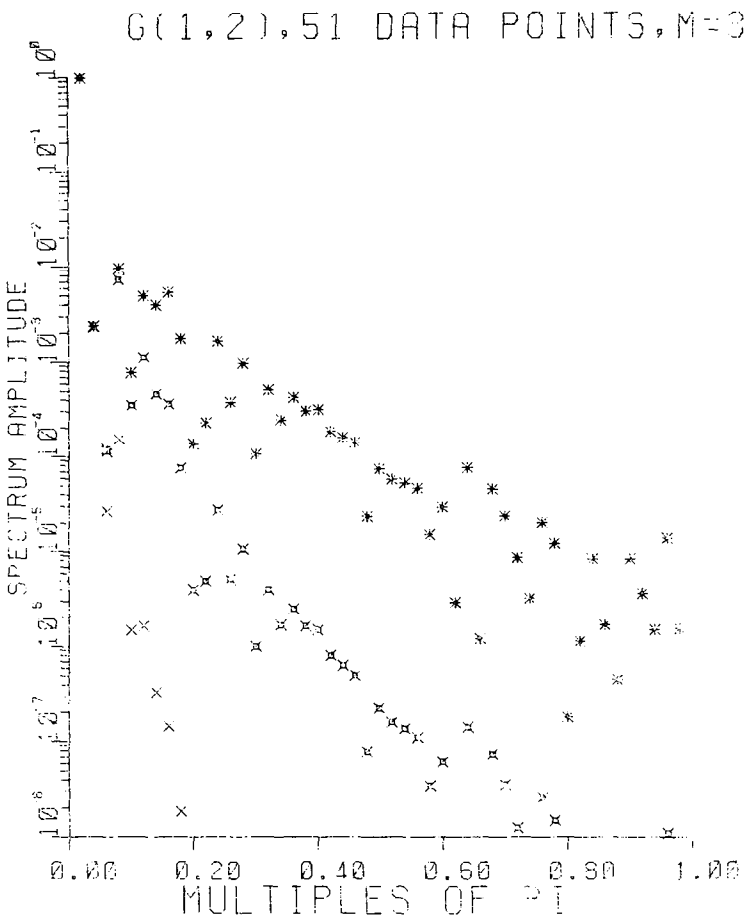

FIG. 1. Increase of spectral smoothing for higher derivatives. Example: g_{12}, $N = 51$, $m = 3$. $(*)$ $n = -1$ $(\tilde{\alpha} = 0.18\text{E}-5)$; $(\square)$ $n = 0$ $(\tilde{\alpha} = 0.96\text{E}-7)$; $(\times)$ $n = 2$ $(\tilde{\alpha} = 0.64\text{E}-9)$.

The same effect can be seen when the order of the smoothing norm is allowed to increase for a required operation (fixed n) to be performed on a given data set (g_{ij}, N). Once again, the optimal value $\tilde{\alpha}$ of the smoothing parameter decreases, but coefficients are smoothed to a greater extent. Figure 2 shows the spectral smoothing required to calculate the second derivative of the function g_{22}, using 51 data points and Sobolev regularizers of increasing order. Although the value of $\tilde{\alpha}$ decreases to almost zero (machine accuracy) for the fourth-order regularizer, increasingly greater spectral smoothing does in fact take place. Changes in the value of $\tilde{\alpha}$ therefore cannot be considered an indication of the extent of smoothing of the data.

5.4. As mentioned in Subsection 5.3, higher-order regularizers, for a given data set (g_{ij}, N) and for a particular required operation on this set (i.e., a given value of n)

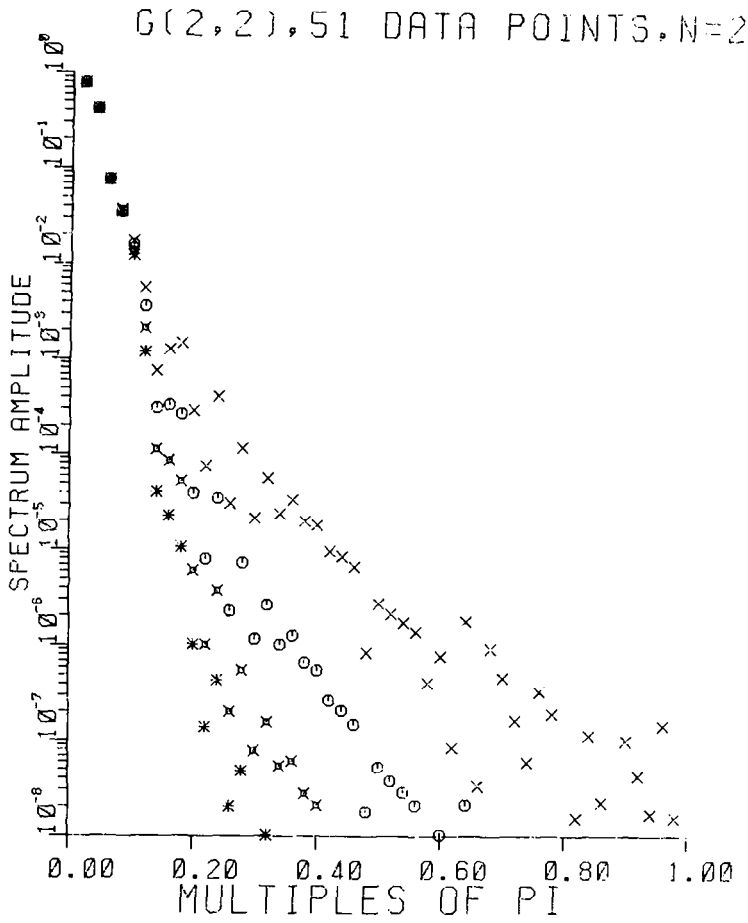

FIG. 2. Increase of spectral smoothing for higher-order norms. Example: g_{22}, $N = 51$, $n = 2$. (×) $m = 1$ ($\tilde{\alpha} = 0.75\,\mathrm{E} - 8$); (Ⓘ) $m = 2$ ($\tilde{\alpha} = 0.63\,\mathrm{E} - 10$); (⊡) $m = 3$ ($\tilde{\alpha} = 0.42\,\mathrm{E} - 12$); (*) $m = 4$ ($\tilde{\alpha} = 0.26\,\mathrm{E} - 14$).

result in greater smoothing of the data. Obviously this can lead to oversmoothing, and the question remains unresolved as to whether the likelihood function Φ can be minimized with respect to m, the order of the Sobolev norm, as well as with respect to the smoothing parameter α; in other words, it is not yet clear whether $\tilde{\Phi}$, the value of the likelihood function for optimal α, can be used as a criterion to determine an optimal smoothing norm.

For the test functions g_{1j}, with bounded derivatives of all orders, the values $\tilde{\Phi}$ as well as E_2 and E_∞ decreased steadily for increasing m, for all data sets and all required operations. For these examples therefore a decrease in $\tilde{\Phi}$ does indicate better approximate solutions. For g_{2j}, however, with $\| g^{(n+m)} \|$ infinite for $(n+m) \geqslant 4$, $\tilde{\Phi}$ did not decrease monotonically with m. Although slight improvements in E_2 and/or

TABLE IV

Changes in $\tilde{\Phi}$ with Increase in m

(g_{ij}, n, N)	m	$\tilde{\Phi} \times 10^{-3}$	E_2	E_∞
$(g_{12}, 2, 101)$	3	−0.284	0.703 E − 2	0.110 E − 1
	4	−0.286	0.597 E − 2	0.968 E − 2
	5	−0.287	0.570 E − 2	0.923 E − 2
	6	−0.288	0.540 E − 2	0.824 E − 2
$(g_{21}, 1, 51)$	1	−0.265	0.217 E − 1	0.471 E − 1
	2	−0.276	0.679 E − 2	0.131 E − 1
	3	−0.269	0.456 E − 2	0.151 E − 1
	4	−0.259	0.554 E − 2	0.259 E − 1
$(g_{23}, 2, 201)$	1	−0.133	0.683 E − 1	0.147
	2	−0.132	0.479 E − 1	0.128
	3	−0.128	0.477 E − 1	0.150
	4	−0.123	0.519 E − 1	0.160

E_∞ occasionally still occurred for $(n + m) \geqslant 4$, this was the exception rather than the rule, and $\tilde{\Phi}$ increased in value for larger values of m. (See Table IV.)

It may be accepted that E_2 and E_∞ are less sensitive to increases in the value of m, and although an increase in $\tilde{\Phi}$ does not necessarily mean an immediate deterioration in the accuracy of the computed solution, it does serve as a warning that higher values of m might not be advisable. The numerical results obtained therefore indicate that the likelihood function (3.7) can be minimized with respect to the order of the Sobolev norm used as regularizer as well as with respect to the smoothing parameter. Furthermore, the optimal regularizer depends on the operation to be performed on the data, i.e., on the kernel of the integral equation (cf. [4]).

ACKNOWLEDGMENTS

I am grateful to Dr. H. Hatze for initiating this study and for the use of his program implementing the A–B algorithm, and to Dr. R. Markham for advice on statistical concepts. Suggestions made by the reviewers led to improved clarity in the presentation of the results.

REFERENCES

1. R. S. ANDERSSEN AND P. BLOOMFIELD, *Numer. Math.* **22** (1974), 157–182.
2. R. S. ANDERSSEN AND P. BLOOMFIELD, *Technometrics* **16** (1974), 69–75.
3. J. CULLUM, *SIAM J. Numer. Anal.* **8** (1971), 254–265.
4. J. CULLUM, *Math. Comput.* **33** (1979), 149–170.
5. H. HATZE, *J. Biomech.* **14** (1981), 13–18.
6. G. R. RICHTER, *SIAM J. Numer. Anal.* **15** (1978), 511–522.
7. A. N. TIHONOV, *Soviet. Math. Dokl.* **4** (1963), 1035–1038.
8. P. WHITTLE, *Skand. Aktuarietidskr.* **35** (1952), 48–60.