JOURNAL OF COMPUTATIONAL PHYSICS 14, 319–340 (1974)

On the Application of Eigenvector Expansions to Numerical Deconvolution*

M. P. EKSTROM AND R. L. RHOADS

Lawrence Livermore Laboratory, University of California, Livermore, California 94550

Received October 20, 1972; Revised November 19, 1973

In this paper a procedure for the numerical solution of convolution-type integral equations is presented. The method uses a spectral representation of a transformed-convolution operator and can be used to solve convolution equations with causal or noncausal kernels. The formalism has the additional property of being both intuitively structured and well-suited for direct numerical computation.

1. INTRODUCTION

A computational problem which arises frequently in mathematical and experimental physics is that of solving linear integral equations of the first kind over a finite interval. Fredholm and Volterra type integral equations of this form are, respectively,

$$r(t) = \int_a^b k(t, \tau)\, x(\tau)\, d\tau, \qquad t \in [a, b] \tag{1}$$

and

$$r(t) = \int_a^t k(t, \tau)\, x(\tau)\, d\tau, \qquad t \in [a, b]. \tag{2}$$

From an applications viewpoint, an important class of such equations involves difference kernels, that is, the kernel $k(t, \tau)$ depends on the difference of its arguments:

$$k(t, \tau) = k(t - \tau). \tag{3}$$

Such integral equations are called convolutions and the process of solving (1) or (2) for $x(t)$ given knowledge of $r(t)$ and $k(t)$ is appropriately called deconvolution. Areas of application where the deconvolution problem occurs are applied optics [1], geophysics [2], communication theory [3], and applied electromagnetics [4].

* This work was performed under the auspices of the U.S. Atomic Energy Commission.

Solving Fredholm and Volterra equations of the first kind is difficult because of the well-known inherent ill-conditioning of the problem. For kernels that occur in physical applications, small variations in $x(t)$ cause practically indistinguishable variations in $r(t)$; however, the converse of this statement is false in rather spectacular ways [since arbitrarily small variations in $r(t)$ can cause arbitrarily large variations in $x(t)$]. Evidently, it is this sensitivity to perturbation or error in $r(t)$ that renders the numerical solution of (1) and (2) nontrivial. As a consequence, the most recent research in this area has been directed toward developing physically acceptable solutions that minimize this sensitivity. Longley [5] proposed replacing the ill-posed problem above with a well-posed one by an ad hoc modification of the kernel function. The more successful solution of Phillips [6] and Twomey [7] is a least-squares procedure which suppresses the large variations in the solution by constraining its second derivative. Baker *et al.* [8] and Hanson [9] suggested solutions based on the smoothing available with a direct spectral decomposition of an integral operator.

In this paper we present an approach to the numerical deconvolution problem which borrows conceptually from the spectral decomposition methods above but which is both applicable to Fredholm and Volterra forms and computationally efficient to apply. The procedure involves using the eigensystem of a transformation of the discrete convolution operator and is beautifully simple in structure. The distinguishing features of the formalism are its ease of computability and its explicit characterization of the solution variations due to the errors or noise in $r(t)$. It is this characterization that allows one to systematically minimize the extraneous variations in constructing an acceptable solution with only crude descriptors of its form.

2. Problem Formulation

The admissible class of functions occurring in (1) and (2) is determined by physical possibility. Consequently, we assume that $x(t)$ and $r(t)$ are square integrable on the whole line and differentiable to all finite orders, and that their Fourier transforms are of bounded support. We call the set of such functions P. Since the kernel characterizes a physical integral operator, it is also assumed to satisfy these conditions. In the Fredholm form of the convolution integral, this kernel is evidently noncausal [that is, $k(t - \tau)$ is nonzero for $t < \tau$]. If the kernel is causal, thus vanishing for $t < \tau$, we can replace the upper limit in (1) by t to give the Volterra form (2).

A variety of quadrature approximations can be used to establish a discrete, finite dimensional analog to these integral operators. Over a finite interval

$t \in [a, b]$, Eq. (1) can be represented by a convolution summation of the form

$$r(n) = \sum_{n'=-A}^{B} K(n - n')\, x(n'). \tag{4}$$

We obtain the sequences $\{r(n)\}$ and $\{x(n)\}$ by sampling the continuous function at equidistant points in the interval $t \in [a, b]$. (For notational convenience, this spacing has been normalized; consequently, A, B, n and n' are integers).

In generating $\{K(n)\}$ the simplest approximation involves a similar sampling of the kernel, that is,

$$K(n - n') = k(n - n'). \tag{5}$$

If the sample spacing is adequately small, consistent with the bandlimits of the continuous functions, the representation (4) using (5) will be exact. If we are not able to control or specify the sample density, a quadrature form which is quite useful and intuitively logical involves picking [5]

$$K(n - n') = \int_{n-n'}^{n-n'+1} k(\lambda)\, d\lambda. \tag{6}$$

It should be clear that by simply reducing the sample spacing, the integral approximation associated with use of either (5) or (6) can be improved to achieve a desired minimum error.

The convolution summation (4) can be written in the following matric notation:

$$\mathbf{r} = \mathbf{K}\mathbf{x},^{1} \tag{7}$$

where

$$\begin{aligned} \mathbf{r} &= \text{response vector} \\ &= \text{col}[r(-A)\, r(-A + 1) \cdots r(B)], \end{aligned} \tag{8}$$

$$\begin{aligned} x &= \text{excitation vector} \\ &= \text{col}[x(-A)\, x(-A + 1) \cdots x(B)], \end{aligned} \tag{9}$$

[1] Upper-case boldface letters denote matrices and lower-case boldface letters denote column vectors.

and

$$\mathbf{K} = \text{kernel matrix}$$

$$\begin{bmatrix} K(0) & K(-1) & \cdots & K(-A-B) \\ K(1) & K(0) & \searrow & \\ \vdots & K(1) & \searrow & \\ K(A+B) & & \searrow & K(0) \end{bmatrix} = [K(i-j)]_{ij=0}^{A+B} \qquad (10)_i$$

Now, in the integral transformation defined by (1) and (2), $x(t)$ is said to be transformed into $r(t)$. The range of this transform is the set of functions R that results from convolution of an admissible kernel with a function $x(t) \in P$. We consider only those kernels such that this transformation is one-to-one and unique. Thus, the inverse of this transform exists and theoretically allows us to associate a given $r(t) \in R$ with $x(t) \in P$. In the finite dimensional characterization, Eq. (7), the matrix $\mathbf{K}^{-1}$ is the discrete analog of the inverse transformation.

While the pairing of $\mathbf{r}$ and $\mathbf{x}$ with $\mathbf{K}^{-1}$ is formally justifiable, it is of limited practical use. This is because, in applications, the accessible response vector contains considerable uncertainty or error. This uncertainty arises from the noise present in all measurement data and, to a lesser extent, the quadrature error in the discretization of (1) and (2). In the presence of this error, (7) becomes

$$\mathbf{r} + \boldsymbol{\delta}\mathbf{r} = \mathbf{K}(\mathbf{x} + \boldsymbol{\delta}\mathbf{x}) \qquad (11)$$

where $\boldsymbol{\delta}\mathbf{r}$ is the error in $\mathbf{r}$ and $\boldsymbol{\delta}\mathbf{x}$ is the error induced in $\mathbf{x}$. Because the error sequences in the response are generally nondifferentiable and wideband in comparison with $\{x(n)\}$ and $\{r(n)\}$, the accessible response is not in the range set, R, of the convolution transformation. Consequently, the excitation $\mathbf{x} + \boldsymbol{\delta}\mathbf{x} \notin P$.

In formally solving (11) we can see how this occurs:

$$\mathbf{x} + \boldsymbol{\delta}\mathbf{x} = \mathbf{K}^{-1}(\mathbf{r} + \boldsymbol{\delta}\mathbf{r}). \qquad (12)$$

It happens in practice that close approximations to $r(t)$ with (7), involving bounded smooth kernels, lead to matrices $\mathbf{K}^{-1}$ with very large numbers. As a result, $\boldsymbol{\delta}\mathbf{x}$ turns out to be extremely large in norm and wildly oscillatory even for very "small" $\boldsymbol{\delta}\mathbf{r}$. To demonstrate this behavior, we consider the kernel sequence and excitation-response pair shown in Figs. 1 and 2, respectively. Note that here we have taken

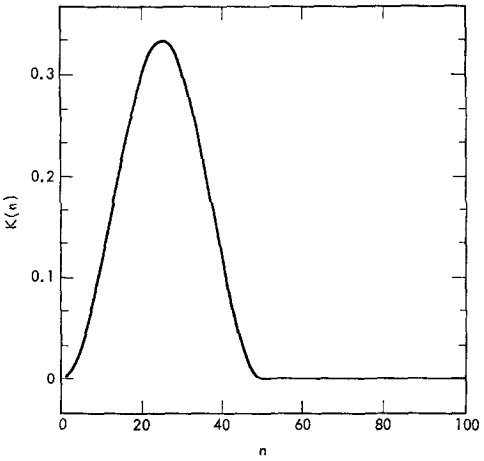

FIG. 1. Kernel sequence $\{K(n)\}$ [obtained from (6)].

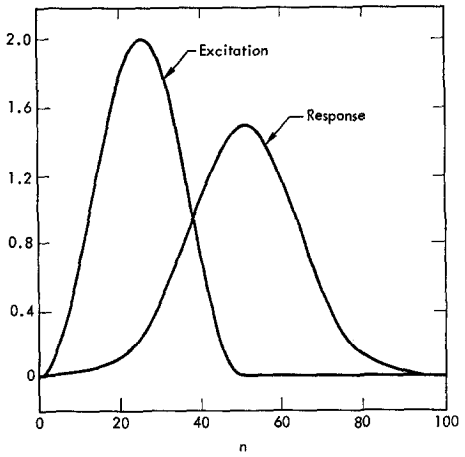

FIG. 2. Excitation $\{x(n)\}$ and response $\{r(n)\}$ [obtained from (7) with $A = 0$ and $B = 100$].

$x(t) = k(t)$. The response vector was generated according to [7] with the elements of $\mathbf{K}$ and $\mathbf{x}$ as shown. (To obtain the *exact* excitation vector given this response, we would, of course, form the matric product $\mathbf{K}^{-1}\mathbf{r}$.) We introduce an error of 10^{-6} in the first element of $\mathbf{r}$, that is,

$$\delta\mathbf{r} = \text{col}[10^{-6}, 0, 0, \ldots, 0],$$

and pair this response to its excitation, $\mathbf{x} + \delta\mathbf{x}$, according to (11). The resulting

excitation is shown in Fig. 3. Judging from its erratic variation and apparent instability, this excitation has few of the fundamental properties we normally associate with physically real solutions, despite the relatively small error in the response vector. This behavior is, of course, quite unsatisfactory and indicates the need for a rational approximation procedure in solving (11).

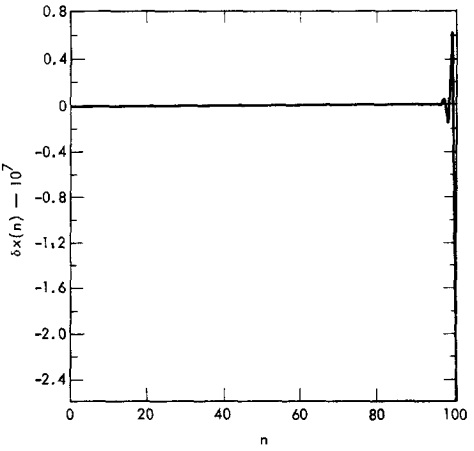

FIG. 3. Excitation error $\{\delta x(n)\}$ due to perturbation of $r(0)$.

3. SOLUTION BY DECOMPOSITION

We apply spectral decomposition techniques in developing a constructive procedure for solving the noisy quadrature system in (11). The general utility of these techniques derives from their classical applicability and convenience in constructing approximate solutions. It is the special structure of the kernel matrix **K** which suggests the spectral decomposition that follows. **K** is a so-called Toeplitz matrix [10] of the general form

$$\mathbf{K} = (K_{i-j})_{ij=0}^{A+B} \tag{13}$$

i.e., it has equal entries on each of its principal diagonals. Because of this structure the system (11) can be also reordered and written in matrix form as

$$\mathbf{r} + \delta\mathbf{r} = \hat{\mathbf{K}}(\hat{\mathbf{x}} + \delta\hat{\mathbf{x}}). \tag{14}$$

Here, the tilde indicates matrix transposition and the caret indicates the exchange of elements in each column about its center. Thus, we have

$$\hat{\tilde{\mathbf{K}}} = \begin{bmatrix} K(-A-B) & \cdots & K(-1) & K(0) \\ \vdots & & & K(1) \\ K(-1) & K(0) & & \\ K(0) & K(1) & \cdots & K(A+B) \end{bmatrix}, \tag{15}$$

and $\hat{\mathbf{x}}$ is just $\mathbf{x}$ turned upside down.

Now, $\hat{\tilde{\mathbf{K}}}$ is a real, symmetric Hankel matrix. As such, it submits to a definitive spectral decomposition, one which is both analytically useful and computationally convenient. Its eigenvalues $\{\lambda_i\}$ are real and, for practical purposes, distinct, and can be ordered as

$$\lambda_0 > \lambda_1 > \cdots > \lambda_{A+B} .$$

The eigenvectors associated with different eigenvalues are linearly independent and orthogonal, and taken together form a basis. Thus, we can express the accessible response in this base as

$$\mathbf{r} + \delta\mathbf{r} = \sum_{i=0}^{A+B} (\mathbf{r} + \delta\mathbf{r}, \mathbf{e}_i)\, \mathbf{e}_i . \tag{16}$$

From (14) we have

$$\hat{\mathbf{x}} + \delta\hat{\mathbf{x}} = (\hat{\tilde{\mathbf{K}}})^{-1} (\mathbf{r} + \delta\mathbf{r}). \tag{17}$$

Plugging (16) into (17) results in

$$\begin{aligned} \hat{\mathbf{x}} + \delta\hat{\mathbf{x}} &= (\hat{\tilde{\mathbf{K}}})^{-1} \sum_{i=0}^{A+B} (\mathbf{r} + \delta\mathbf{r}, \mathbf{e}_i)\, \mathbf{e}_i \\ &= \sum_{i=0}^{A+B} \frac{1}{\lambda_i} (\mathbf{r} + \delta\mathbf{r}, \mathbf{e}_i)\, \mathbf{e}_i , \end{aligned} \tag{18}$$

where this last relation follows from the interdependence of the eigensystems of $\tilde{\mathbf{K}}$ and its inverse. Consequently, the solution vector can be written as

$$\mathbf{x} + \delta\mathbf{x} = \sum_{i=0}^{A+B} \beta_i \hat{\mathbf{e}}_i , \tag{19}$$

where

$$\begin{aligned}\beta_i &= \frac{1}{\lambda_i}(\mathbf{r} + \delta\mathbf{r}, \mathbf{e}_i) \\ &= \frac{1}{\lambda_i}(\mathbf{r}, \mathbf{e}_i)\left[1 + \frac{(\delta\mathbf{r}, \mathbf{e}_i)}{(\mathbf{r}, e_i)}\right].\end{aligned} \tag{20}$$

Evidently, as a formal solution to Eq. (11), Eq. (19) suffers from the same practical difficulties as Eq. (12). In (19) and (20) however, we have the effect of the response $\delta\mathbf{r}$ explicitly characterized in a form that suggests a rational approach to its minimization. It is evident that the error in each term of the eigenvector expansion of $\mathbf{x} + \delta\mathbf{x}$ is completely accounted for in (20) by a noise-to-signal ratio $(\delta\mathbf{r}, \mathbf{e}_i)/(\mathbf{r}, \mathbf{e}_i)$. This error will be objectionable if this ratio is on the order of 1.0 or larger, and/or $|\lambda_i|$ is small. A similar spectral dependency can also be inferred from the cumulative error $\|\delta\mathbf{x}\|$ since it can be shown that this norm can be bounded (sometimes tightly) by

$$\|\delta\mathbf{x}\| \leqslant \frac{|\lambda|_{\max}}{|\lambda|_{\min}} \|\delta\mathbf{r}\|. \tag{21}$$

Thus, it follows that with the above decomposition the main components of $\delta\mathbf{x}$ can, to some extent, be spectrally isolated.

These considerations lead to the following procedure for developing physically acceptable solutions of (11). We form the approximate solution $\mathbf{x}_a$ by a generalized Fourier expansion using the basis $\{\mathbf{e}_i\}$. The coefficients of this expansion are obtained by a weighting of the Fourier coefficients in (20). Specifically, we have

$$\mathbf{x}_a = \sum_{i=0}^{A+B} W(\lambda_i)\, \beta_i \hat{\mathbf{e}}_i \tag{22}$$

where $\{W(\lambda_i)\}$ is a weight or penalty sequence. This weighting is used to control the behavior of the approximation by minimizing the effects of those terms in the expansion which contribute significantly to $\delta\mathbf{x}$. Frequently, in applications, these terms can be identified by analyzing the measurement system and/or the physics involved in the experiment. In any event, the weighting sequence must be specified a priori and in the next section we present some candidate weighting schemes that we have found effective for this purpose.

The numerical calculations required in construcing $\mathbf{x}_a$ in (22) are quite straightforward to implement. The most difficult problem involves computing the eigensystem of $\hat{\mathbf{K}}$. We use an algorithm written by J. H. Wilkinson and others, which was specifically designed for handling real-symmetric matrices. It involves a reduction of $\hat{\mathbf{K}}$ to a tridiagonal symmetric form using Householder's method [11]

and the subsequent determination of its eigenvalues and eigenvectors using QL transformations [12]. We have found this algorithm suitable for handling quite large matrices ($A + B \approx 400$). Once the eigensystem is computed, of course, it can be stored and used for all deconvolution problems involving that particular kernel.

As we mentioned previously, our decomposition method is similar in spirit to that of Baker *et al.* [8] and Hanson [9]. In contrast to this first method, however, our formalism is directly applicable to Volterra and nonsymmetric Fredholm forms. As we mentioned previously, solution of Volterra equations (i.e., causal kernels) represents a particularly important class of problems. The basis of contrast with the second method is fundamentally computational. Our decomposition technique involves better conditioned eigensystem computations and requires $(A + B + 1)^2$ fewer storage elements. For large dimension problems, these may be extremely important considerations.

4. Weighting Schemes and Examples of Application

It is an appropriate choice of the weighting sequence in (22) that determines the usefulness of $\mathbf{x}_a$. In practice, the particular choice of $\{W(\lambda_i)\}$ is generally motivated by physical considerations. These have to do with the overall confidence one has in the spectral components of $\mathbf{r} + \boldsymbol{\delta}\mathbf{r}$ and the physical acceptability of the constructed solution. Both considerations can be accommodated in a relatively straightforward manner by picking the elements of the weighting sequence $\{W_i\}$ to be of the general form

$$W_i = \frac{|\lambda_i|}{|\lambda_i| + \epsilon_i}. \tag{23}$$

This form allows weighting to a wide variety of criteria by virtue of its ease of interpretability.

Below, we present several numerical examples that demonstrate the methodology used in selecting this weighting and the application of our formalism for solving deconvolution problems. We compared our solutions with those available from other integral equation solvers that employ essentially equivalent a priori knowledge of the solution. It is important to note that, for the class of problems normally encountered in experimental physics (and for the numerical examples considered below), the dominate errors in the solutions arise from measurement errors in $\mathbf{r}$ and *not* from inaccuracy in the quadrature approximation of (2). This property is a determining factor in our solution formalism and is in contrast with much of the existing literature on the solution of linear integral equations (see, for example, [13–15]).

Example 1

We first illustrate the application of our deconvolution method by considering the Volterra equation formulation of the classic problem introduced by Phillips [6] and used in our previous example. In this example, we want to solve (2) with

$$x(t) = 6 \times k(t) = \begin{cases} \left[1 + \cos \dfrac{\pi(t-3)}{3}\right] & 0 \leqslant t \leqslant 6, \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

where $k(t, \tau)$ is a difference kernel. The kernel and excitation sequences are generated according to (4) and (6) with $A = 0$, $B = 100$, and the sampling interval equal to 0.12. As we saw in Fig. 3, the exact solution this system is quite sensitive to the presence of errors in $\mathbf{r}$. Following Phillips' example, we generate an error sequence $\boldsymbol{\delta}\mathbf{r}$ by quantizing or rounding-off the components of $\mathbf{r}$ such that the maximum value of $\boldsymbol{\delta}\mathbf{r}$ would be ± 0.005. This error is shown in Fig. 4.

With $\mathbf{r} + \boldsymbol{\delta}\mathbf{r}$ in hand, our first step in solving for $\mathbf{x}_a$ involves computing the eigensytem of $\hat{\tilde{\mathbf{K}}}$. The computed eigenvalues are shown in Fig. 5. Now, in this case, the ratio $|\lambda|_{\max}/|\lambda|_{\min}$, which is a measure of the ill-conditioning or sensitivity of the solution to response errors, is on the order of 5×10^{20}.

In picking the weighting sequence for this example, we assume little a priori knowledge of the solution, only that is of finite norm. This involves letting

$$\epsilon_i = \epsilon. \tag{25}$$

It is easy to see that this corresponds to picking the weights to discriminate against

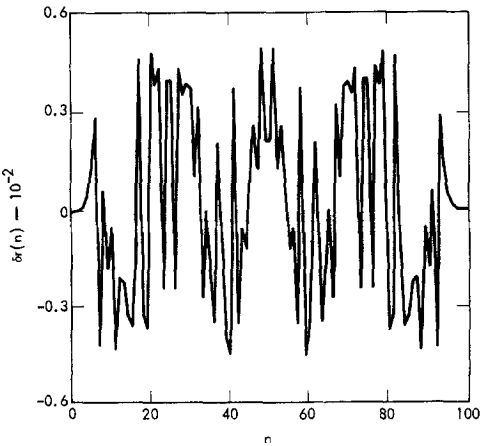

FIG. 4. Round-off error $\{\delta r(n)\}$ in the response sequence.

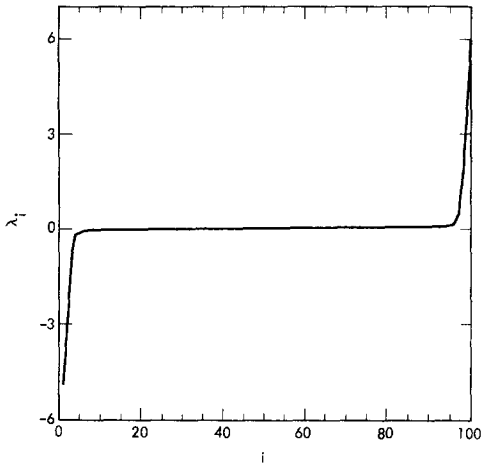

FIG. 5. Eigenvalues of $\tilde{\tilde{\mathbf{K}}}$ for the text Example 1.

those components associated with the smaller λ_i. For "large" $|\lambda_i|$, $W_i \approx 1$, while for "small" $|\lambda_i|$, $W_i \approx 0$. We decide in effect what is large and small by assigning the constant, ϵ. This weighting is useful in many experimental applications where a band-limited response is measured in the presence of wideband or white noise.

Using (23) and (25), we tried various values of ϵ over the range 0–10, forming the solution according to (22). The error norm $\| \mathbf{x} - \mathbf{x}_a \|$ was then evaluated for each ϵ. The variation of this error norm versus ϵ is shown in Fig. 6. It is characteristic

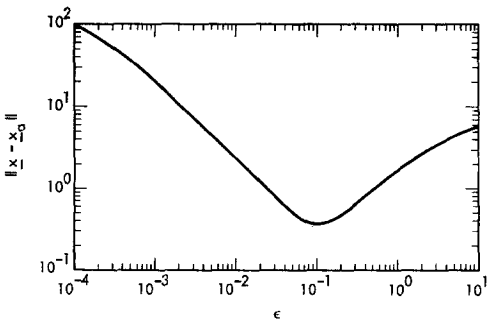

FIG. 6. Norm of the solution error $\| \mathbf{x} - \mathbf{x}_a \|$ for the case of ϵ_i = constant.

of the effect of $\{W_i\}$ on $\mathbf{x}_a$. For small ϵ (all $W_i \approx 1$), the error in $\mathbf{x}_a$ is quite large because of the noise in the response vector. Conversely, for large ϵ (all $W_i \approx 0$), the large error is a result of the loss of components in $\mathbf{x}_a$ associated with the larger $|\lambda_i|$. Between these two extremes lies the optimum choice of ϵ. The sequenc $\{W_i\}$

and the approximate solution $\mathbf{x}_a$ corresponding to the best ϵ are given in Figs. 7 and 8, respectively. Apparently, this $\mathbf{x}_a$ is a close representation of the true excitation with its important characteristics (peak value, full-width-at-half-maximum, and general bell shape) preserved in the solution.

We now compare this result with a solution obtained by one of the methods proposed for numerically solving analytical integral equations (i.e., those that are exactly known). This method is due to Jones [13]. It is a direct, iterative solution

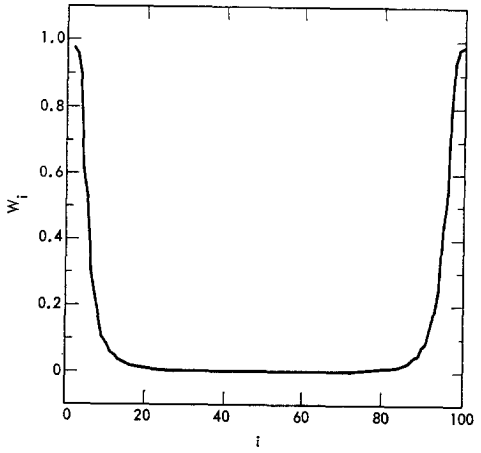

FIG. 7. Optimum weighting sequence $\{W_i\}$ for $\epsilon = 0.1$.

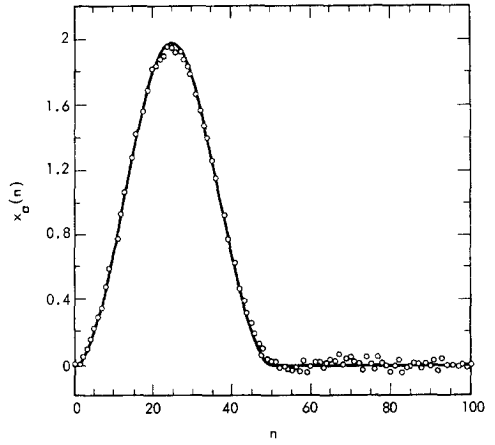

FIG. 8. Approximate solution $\{x_a(n)\}$ corresponding to the weights of Fig. 7 (solid line is $\{x(n)\}$).

of the linear system [11], particularly well-suited for numerical implementation. The solution so obtained is shown in Fig. 9, and it is easy to see that it is a completely unacceptable result. The error norm $\| \mathbf{x} - \mathbf{x}_a \|$ is on the order of 2×10^3. Interestingly, improvement of the approximation error either by decreasing the sample spacing or by increasing the order of the quadrature form leads to a more unsatisfactory solution (this is due to the increased ill-conditioning of the linear system with no corresponding reduction of experimental error in $\mathbf{r}$).

Example 2

A desirable feature in any deconvolution procedure should be its ability to use additional a priori information in improving the constructed solution. In this

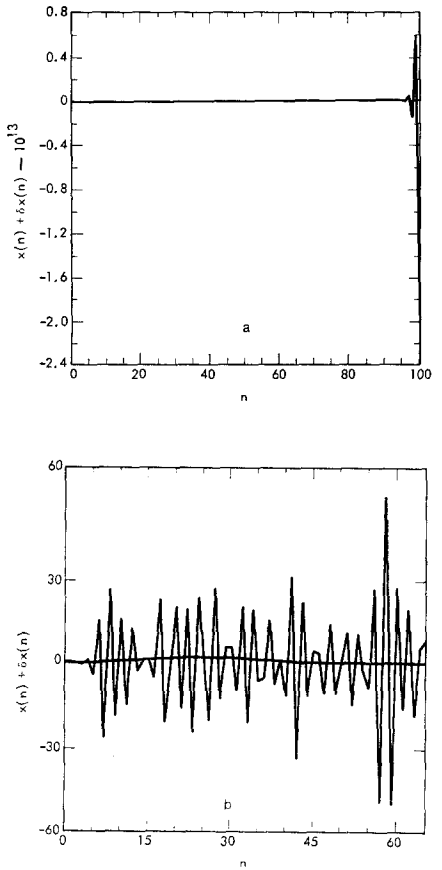

FIG. 9. Solution $\{x(n) + \delta x(n)\}$ obtained with a direct solution: (a) total sequence; (b) first 65 points (solid line is $\{x(n)\}$).

example, we consider the problem immediately above and show that, if some measure of the smoothness of $\mathbf{x}$ is known, an alternative weighting procedure can be used, which eliminates most of the random fluctuations seen in Fig. 8.

This particular weighting involves discriminating against components associated with eigenvectors having high variation between elements. We do this by letting

$$\epsilon_i = \epsilon \sum_{j=0}^{N-1} (e_{j+1}^i - e_j^i)^2, \tag{26}$$

where e_j^i is the jth component of the ith eigenvector. Again, we evaluate this weighting for a range of ϵ and show the error behavior in Fig. 10. To the extent that the smaller eigenvalues of $\hat{\tilde{\mathbf{K}}}$ correspond to eigenvectors with higher variations, this weighting will be somewhat similar to the previous case, although, as is seen in Fig. 11 [the optimum $\{W_i\}$], it is noticeably more discriminate. The solution obtained with the $\{W_i\}$ of Fig. 11 is a remarkably faithful replica of $\mathbf{x}$ and is shown in Fig. 12. An examination of $\| \mathbf{x} - \mathbf{x}_a \|$ indicates a 50% reduction in the solution error norm over that obtained previously. The elements of $\delta\mathbf{x}$ are now on the same order of magnitude as the quantization errors in $\mathbf{r}$.

A method of Schmaedeke [16] deals with the solution of (2) in a more general, separable Hilbert space setting. For comparison, we treated this numerical example using his method. Schmaedeke's solution form that most closely corresponds to the a priori information used here is given (in our notation) by

$$\mathbf{x}_s = (\tilde{\mathbf{K}}\mathbf{K} + \alpha^2\mathbf{D})^{-1}\tilde{\mathbf{K}}\mathbf{r}, \tag{27}$$

where $\mathbf{D}$ is the product of the adjoint of the first difference operator and the first difference operator. The parameter α^2 is initially unknown but is iteratively adjusted

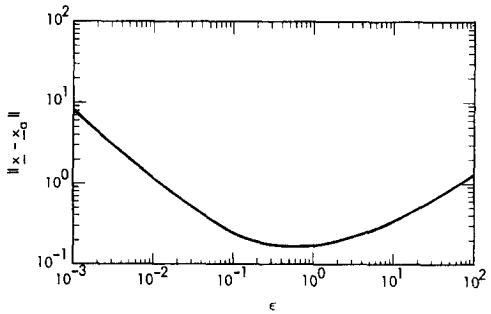

FIG. 10. Norm of solution error $\| \mathbf{x} - \mathbf{x}_a \|$ for case of

$$\epsilon_i = \epsilon \sum_{j=0}^{100} (e_{j+1}^i - e_j^i)^2.$$

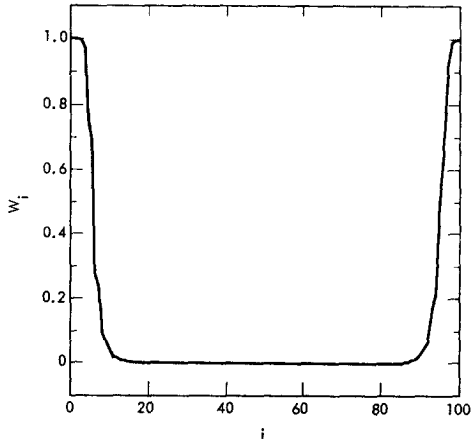

FIG. 11. Optimum weighting sequence $\{W_i\}$ for $\epsilon = 0.6$.

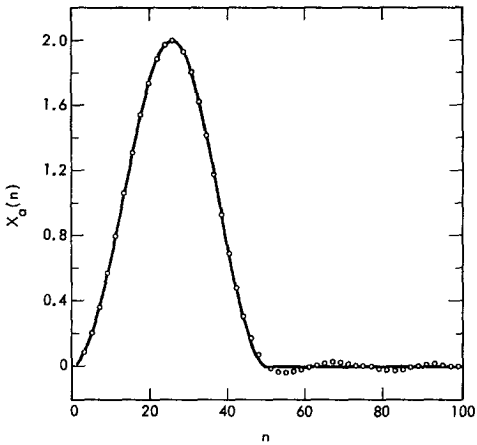

FIG. 12. Approximate solution $\{x_a(n)\}$ corresponding to the weights of Fig. 10 (solid line is $\{x(n)\}$).

to ensure that the norm of the first variation of $\mathbf{x}_s$ satisfies a prescribed value. Several iterations were required to find the appropriate α^2. For this solution we had $\| \mathbf{x} - \mathbf{x}_s \| \approx 2 \times 10^{-4}$, which is essentially equivalent to our result above.

While the error performance measure is directly comparable, the complexity of computation involved with implementing the two schemes is considerably different. Schmaedeke's method requires inversion of an $A + B + 1$ dimension matrix for every iteration in determining the correct α^2. In contrast, we need only

compute the eigensystem of $\hat{\tilde{\mathbf{K}}}$ and the appropriate inner products once in finding the optimum ϵ. In addition, an examination of the weighting indicates that the calculations in forming $\mathbf{x}_a$ can be further reduced by simply eliminating those terms in (22) for which the weight is "essentially" zero. This set of terms can be determined by a direct examination of $\{W_i\}$ and/or the eigensystem of $\tilde{\mathbf{K}}$. Thus, it is clear that computationally the eigenvector expansion is much simpler to use.

Example 3

In this example we consider a problem frequently encountered in computational physics, that of differentiating a finite set of experimental data. In the continuous case, differentiation of a function $x(t)$ is equivalent to solving a convolution equation of the form (2) (with the assumption that $r(a) = 0$), where the kernel $k(t)$ is the unit step function. As in the above examples, instability and high sensitivity to error characterize the standard differentiation procedures for exact data when used with experimental data [17].

In applying our solution formalism to the differentiation problem, we take $x(t)$ as given in (24) and discretize the integral equation as in Example 1, above. The response vector $\mathbf{r}$ is formed according to (7) and augmented by an error vector $\delta\mathbf{r}$. Here, the error sequence is a white noise process, each element of which is a realization of a random variable uniformly distributed on the interval $[-0.3, +0.3]$. The finite data set $\mathbf{r} + \delta\mathbf{r}$ which we wish to differentiate is shown in Fig. 13. Qualitatively, we would describe this data as being moderately noisy.

In applying the eigenvector solution formalism to this differentiation problem,

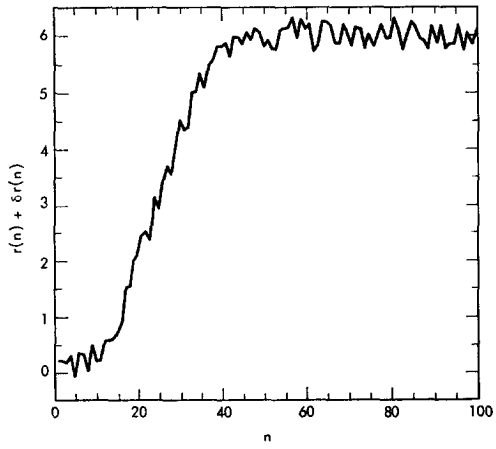

FIG. 13. Response sequence, $\mathbf{r} + \delta\mathbf{r}$, to be differentiated.

we again assume some a priori knowledge of the smoothness of the derivative. In this case we pick

$$\epsilon_i = \epsilon \sum_{j=0}^{N-1} (e_{j+1}^i - 2e_j^i + e_{j-1}^i)^2, \tag{28}$$

thereby minimizing the contribution to $\mathbf{x}_a$ in (22) of those components associated with eigenvectors having high second differences. An examination of the spectrum of the unit step kernel (see Fig. 14) indicates the ratio $|\lambda|_{\max}/|\lambda|_{\min}$ is on the order of 130. The variation of the error norm as a function of ϵ is shown in Fig. 15, and in Fig. 16 we have the optimum weighting sequence. The solution constructed with this weighting is given in Fig. 17 is a remarkably faithful replica of the true

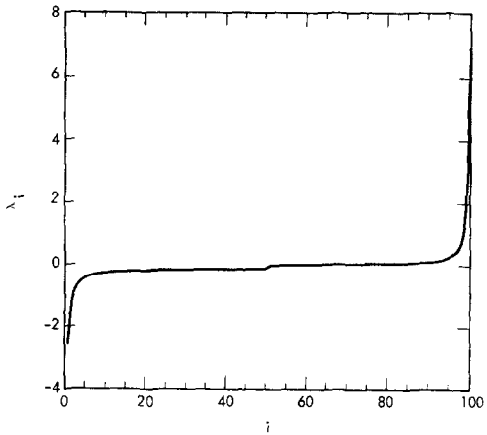

FIG. 14. Eigenvalues of K for the text Example 3.

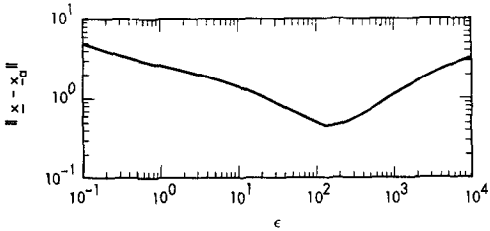

FIG. 15. Norm of solution error $\| \mathbf{x} - \mathbf{x}_a \|$ for the case of

$$\epsilon_i = \epsilon \sum_{j=0}^{100} (e_{j+1}^i - 2e_j^i + e_{j-1}^i)^2 .$$

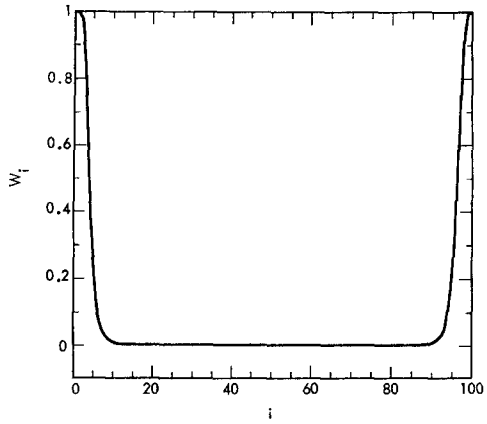

FIG. 16. Optimum weighting sequence for $\epsilon = 150$.

derivative. (Numerical experiments with this example, using Gaussian-distributed white noise of equal variance, yield sensibly equivalent results.)

For comparative purposes we now apply a more classical differentiation scheme to this example. This scheme is characteristic of the class of integral equation solvers in which the inversion of the equation can be performed analytically in closed form [18, 19]. These inversions are not generally without anomaly because of their noise amplification property. Thus, while they are suitable for exact problems, they are typically not reliable for handling experimental data. One approach to adapting this technique to nonexact problems involves approximating the data by

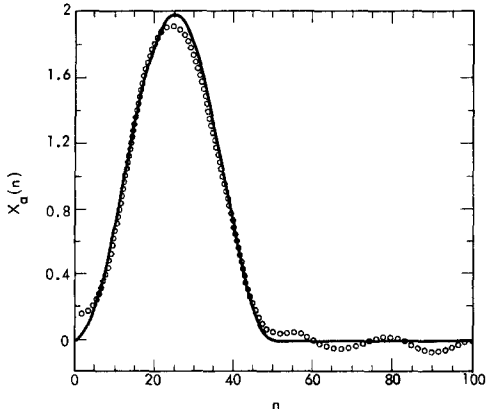

FIG. 17. Approximate derivative $\{x_a(n)\}$ corresponding to the weights of Fig. 16 (solid line is $\{x(n)\}$).

a set of smooth analytic functions and applying the inversion to this representation.

The differentiation method that we now demonstrate is structured along these lines and is similar to one proposed by Savitzky and Golay [20]. The data are first approximated in a least squares sense over a sliding subinterval by a fixed-order polynomial. This polynomial is differentiated (the inversion) at the center point of the subinterval, to give the estimated value of the derivative at that point. The subinterval is then advanced one point and the procedure repeated until the entire interval has been transversed. The polynomial approximation is performed by using orthogonal polynomials, and its "smoothing" properties are controlled by varying the length of the subinterval (i.e., fit interval).

Applying this technique and adjusting for the optimum fit interval, we obtained the estimated derivative in Fig. 18. Its error $\| x - x_d \| \approx 1.2$ is to be compared with the estimated derivative of Fig. 17, whose error norm is $\sim$0.4. In this case superiority of the eigenvector solution is clearly evident.

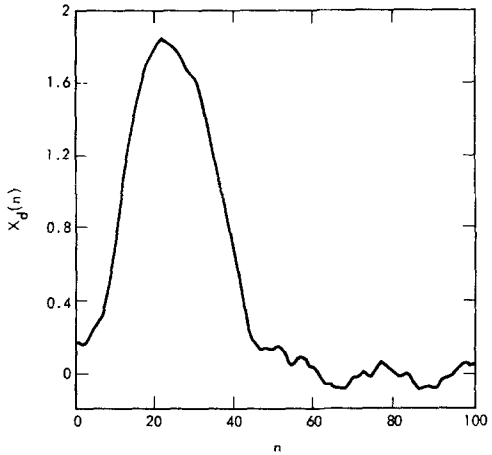

FIG. 18. Approximate derivative $\{x_d(n)\}$ obtained with a Savitzky/Golay-type algorithm.

4. CONCLUSION

We have shown in the above examples that the deconvolution procedure (22) leads to useful results with an appropriate choice of ϵ in the above weighting schemes. It remains to be shown whether a near-optimum choice of ϵ can be made when $\mathbf{x}$ is unknown. In this regard, our experience in applying the method to a variety of problems may be of interest. We have found that the error behavior in Figs. 6, 9, and 14 is typical in that the minimum of $\| \mathbf{x} - \mathbf{x}_a \|$ is relatively

insensitive to variations in ϵ. Thus, while it is clearly important to obtain a "close" estimate of the optimum ϵ, the degree of closeness is not crucial. Further, our earlier remarks on the behavior of $\| \mathbf{x} - \mathbf{x}_a \|$ suggest a rationale we have found satisfactory for picking a sufficiently close ϵ. We noted that as ϵ increased past the optimum value we eventually began eliminating relevant components of the true solution. This effect is essentially an over-smoothing and leads to the observed upturn in $\| \mathbf{x} - \mathbf{x}_a \|$. Because these components are associated with the larger eigenvalues, it follows that this effect should be likewise evident in $\| \mathbf{r} + \delta\mathbf{r} - \mathbf{x}_a \|$. In Fig. 19

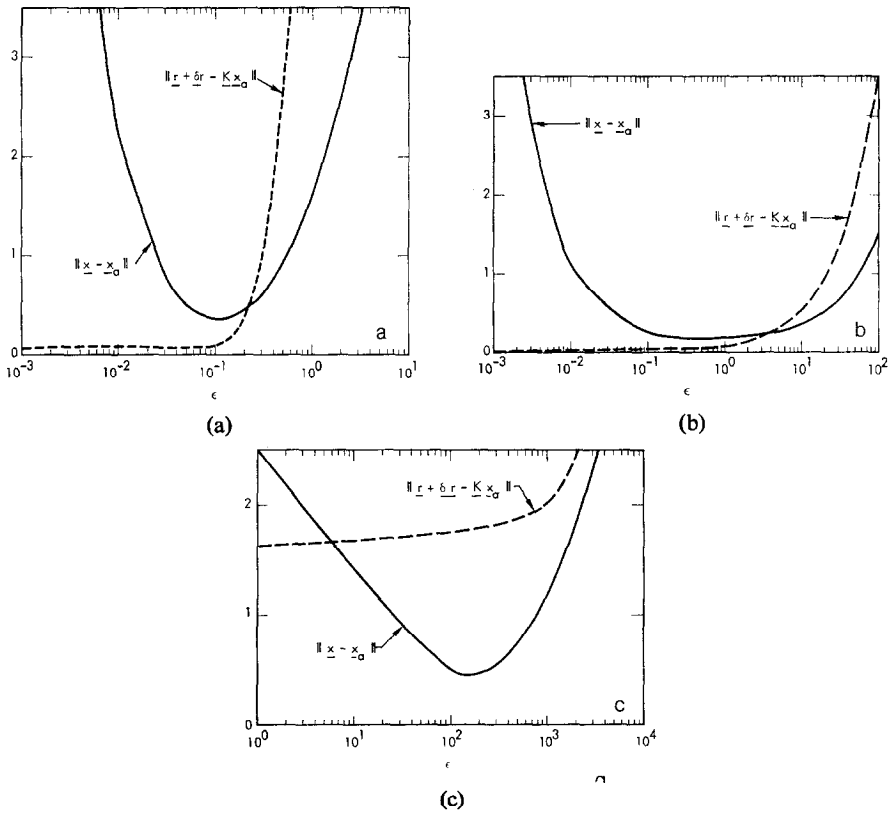

FIG. 19. Variation of solution error norm and $\| \mathbf{r} + \delta\mathbf{r} - \mathbf{K}\mathbf{x}_a \|$: (a) Example 1, where ϵ_i = constant = ϵ; (b) Example 2 where

$$\epsilon_i = \epsilon \sum_{j=0}^{100} (e_{j+1}^i - e_j^i)^2; \text{ and}$$

(c) Example 3, where

$$\epsilon_i = \epsilon \sum_{j=0}^{100} (e_{j+1}^i - 2e_j^i + e_{j-1}^i)^2.$$

we show the solution and response errors associated with $\mathbf{x}_a$ for the three examples and see that this is indeed the case. For small ϵ, $\| \mathbf{r} + \delta\mathbf{r} - \equiv\mathbf{x}_a \|$ is quite small, increasing only gradually to the region of optimum ϵ. As ϵ increases further, a noticeable increase in this norm occurs, almost concurrent with the increase in $\| \mathbf{x} - \mathbf{x}_a \|$.

This behavior can be deduced more formally using norm arguments and has been experimentally verified by the authors in applications involving both correlated and uncorrelated error processes. This is indeed a fortunate circumstance, since it allows one to pick a reasonably good estimate of the optimum ϵ by simply searching the region immediately preceding the upturn. Some judgment and experience may be helpful in settling on a final value for ϵ; however, the selection procedure involves only that information which is available in actual applications.

We note in conclusion that a myriad of useful weighting schemes is of course available. We have simply presented three to demonstrate the rationale of their selection. Common characteristics of other schemes should be their capacity to incorporate practical judgement into the solution formalism and their adaptability to the physical problem at hand.

ACKNOWLEDGMENT

The author wishes to thank R. N. Castleton for his programming support and his assistance with the numerical examples presented in this paper.

REFERENCES

1. C. W. BARNES, *J. Opt. Soc. Amer.* **56** (1966), 575–578.
2. G. KUNETZ AND J. M. FOURMANN, *Geophysics* **33** (1968), 412–423.
3. D. SLEPIAN AND T. T. KADOTA, *SIAM J. Appl. Math.* **17** (1969), 1102–1117.
4. K. K. MEI, *IEEE Trans. Antennas. Propagat.* **13** (1965), 374–378.
5. H. J. LONGLEY, "Numerical Solutions and Applications of the Fold Integral," Los Alamos Scientific Laboratory, Rept. LA-2729 (1962).
6. D. L. PHILLIPS, *J. Assoc. Comput. Mach.* **9** (1962), 84–97.
7. S. TWOMEY, *J. Franklin Inst.* **279** (1965), 95–109.
8. C. T. H. BAKER, L. FOX, D. F. MAYERS, AND K. WRIGHT, *Comput. J.* **7** (1964), 141–148.
9. R. J. HANSON, *SIAM J. Numer. Anal.* **8** (1971), 616–622.
10. U. GRENANDER AND G. SZEGÖ, "Toeplitz Forms and Their Applications," University of California Press, Berkeley, 1958.
11. R. S. MARTIN, C. REINSCH, AND J. H. WILKINSON, *Numer. Math.* **11** (1968), 181–195.
12. H. BOWDLER, R. S. MARTIN, C. REINSCH, AND J. H. WILKINSON, *Numer. Math.* **11** (1968), 293–306.
13. J. G. JONES, *Math. Comp.* **15** (1961), 131–142.
14. L. B. RALL, *SIAM Rev.* **7** (1965), 55–64.

15. P. LINZ, *Comput. J.* **12** (1969), 393–397.
16. W. W. SCHMAEDEKE, *J. Math. Anal. Appl.* **23** (1968), 604–613.
17. R. S. ANDERSSEN AND P. BLOOMFIELD, "On the Numerical Differentiation of Data," Department of Statistics, Tech. Rept. 13, Princeton University, April 1972.
18. M. P. FREEMAN AND S. KATZ, *J. Opt. Soc. Amer.* **50** (1969), 826–830.
19. G. N. MINERBO AND M. E. LEVY, *SIAM J. Numer. Anal.* **6** (1969), 598–616.
20. A. SAVITZKY AND J. E. GOLAY, *Anal. Chem.* **36** (1964), 1627–1639.