JOURNAL OF COMPUTATIONAL PHYSICS 3, 486–493 (1969)

A Certain Unfolding Problem*

GEORGE A. BAKER, JR.

Applied Mathematics Department, Brookhaven National Laboratory, Upton, New York 11973

Received September 13, 1968

ABSTRACT

The problem of unfolding a distribution which is a composite of several unknown but identical distributions is considered. A numerical method is developed, based on a smoothing criterion, which, when applied to a sample set of experimental data, yields results in good agreement with previous hand-graphical procedures.

1. INTRODUCTION AND SUMMARY

In this paper we consider the problem of unfolding a distribution which is a composite of several unknown, but identical distributions. The experimental data, y, is of the form

$$y(x) = \int d\xi\, f(x - \xi) \sum_{i=1}^{n} \alpha_i \delta(\xi - \tau_i)$$

with f unknown except for being "smooth". The problem is to find the (α_i/α_1) and $(\tau_i - \tau_1)$. We set up a smoothing criterion,

$$\delta \left\{ \int (f')^2 \, c(x)\, dx \right\} = 0$$

and, subject to fitting the experimental data, derive from it a system of equations. These equations are of the form

$$\sum_{j=1}^{N} (\mathbf{A}_{ij} + \mathbf{B}_{ij}) f_j = u_i$$

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

where **A** and **B** are separately factorable into the product of an upper and a lower triangular matrix, but the sum is not transparently so factorable. A factorization, pseudotime method is introduced, following the Baker–Oliphant [1, 2, 3] method for multi-dimensional heat flow problems where a similar structure arises. The application of this method to a sample problem results in an error reduction of a factor of 100 per 7 to 15 "major cycle" iterations. In the sample problem considered, smoothing the curve was significant and the results were in good agreement with hand-graphical procedures used heretofore.

2. Description of the Problem

Although the procedures we develop are applicable to a wide class of problems, we shall, for the sake of clarity concentrate on the following unfolding problem which arises in nuclear physics. Samples of material are irradiated and then allowed to decay through the process of internal conversion, i.e. the energy of decay is used to eject an electron from the atom which surrounds the decaying nucleus. The ejected electrons are then passed through a spectrometer, which analyzes them according to energy. Electrons starting in the same atomic subshell will have very close to the same energy. If perfect experimental equipment were available, one would expect to see a family (3 in the case we will consider) of very narrow, well defined lines. Due, however, to imperfections in the apparatus, each well defined original line produces an observed line shape (vs. energy) which characteristically drops fairly sharply to zero on the high energy side, but which has a long decaying tail on the low energy side. This shape is not known theoretically but is a complicated functions of the experimental apparatus used. What is known about it is, that it is the same, to all intents and purposes, for all (three) the lines of a family because they lie so close together in energy. The line shape will vary only slowly as a function of the energy of the line.

The quantities of interest are the relative strengths (total intensities) and positions (energies) of the lines. We may break the data into three regions in energy, where one, two, and three of the lines respectively make a nonzero contribution to the experimental counting rate.

If $y(x)$ is the observed intensity at energy (or actually momentum which is correlated to energy in a one–to–one fashion) x and $f(x)$ is the true shape function for the line with the highest momentum, then we have

$$\begin{aligned} &\text{region I} && y(x) = f(x) + \epsilon(x) \\ &\text{region II} && y(x) = f(x) + \alpha f(x + T_1) + \epsilon(x) \\ &\text{region III} && y(x) = f(x) + \alpha f(x + T_1) + \beta f(x + T_2) + \epsilon(x). \end{aligned} \tag{2.1}$$

where T_1 and T_2 are the shift in momentum between the positions of the first and second, and the first and third lines respectively. $\epsilon(x)$ is the experimental error. The multipliers α and β are the ratio of intensities between the second and first, and the third lines respectively. Since the data is taken at discrete points we may more compactly and conveniently represent (2.1) as

$$Y_i = \mathbf{M}_{ij} f_j + \epsilon_i \tag{2.2}$$

where $\mathbf{M}_{ij}$ is an upper triangular matrix. Summation over a repeated index is implied. When the translations T_1 and T_2 are not an integral number of data points, which in general they will not be, then an interpolation is implied to find $f(x + T_1)$ and $f(x + T_2)$. We have used, in practice, 4-point interpolation.

The problem then is, from an experimental set of $y(x)$, to determine α, β, T_1 and T_2.

3. Smoothing Criterion

In order to make the problem set forth in the previous section meaningful, it is necessary to introduce a smoothing criterion. Various possible criteria have been proposed ([4], [5], [6]). Podolsky and Denman [7] have proved that if the smoothing is to be scale free and imposed by a condition

$$\delta \left\{ \int_a^b w(x, f', f'')\, dx \right\} = 0 \tag{3.1}$$

then w must be a homogeneous function of f' and f''. We prefer the following criterion

$$\delta \left\{ \int_a^b (f')^2\, c(x)\, dx \right\} = 0 \tag{3.2}$$

where $c(x) \geqslant 0$ because we feel intuitively that the shortest curve would be the best one and the length of a curve is

$$\int_{x=a}^{b} ds = \int_a^b [1 + (df/dx)^2]^{1/2}\, dx \tag{3.3}$$

The shortest-curve condition is not scale free however. To make it scale free [7] we observe if we introduce a scale factor ϵ, which we think of as making $\max | \epsilon\, df/dx |$ small compared to unity, then (3.3) becomes

$$\int_a^b \left[1 + \frac{1}{2}(\epsilon\, df/dx)^2 + \cdots\right] dx = (b - a) + \frac{1}{2}\epsilon^2 \int_a^b \left(\frac{df}{dx}\right)^2 dx + \cdots \tag{3.4}$$

As $(b - a)$ has a fixed value, we may disregard it for variational purposes. The function $c(x)$ serves to generalize the scale of the independent variable.

In addition to the requirement that $f(x)$ be smooth, (3.2), we shall also require that the experimental data fit (2.2) within experimental error. To this end we impose

$$\sum_i (y_i - \mathbf{M}_{ij} f_j)^2 w_i \leqslant S_0 \tag{3.5}$$

a least square procedure, where the weights are one over the experimental counts (before the background is subtracted out). The expected value of (3.5) is [8] N, the number of data points.

Consequently, we now impose the requirements that the discrete analogue of (3.2) be a minimum with respect to f_i, α, β, T_1 and T_2, subject to (3.5). Using the standard method of Lagrange multipliers [9] we obtain the equations

$$\delta(Q)/\delta f_i = 0 \tag{3.6}$$

where

$$Q = \left[\sum_{i=1}^{N-1} \frac{(f_{i+1} - f_i)^2}{x_{i+1} - x_i} c(x_i) + \lambda \sum_{i=1}^{N} (y_i - \mathbf{M}_{ij} f_j)^2 w_i\right] \tag{3.7}$$

or writing out (3.6)

$$\frac{f_{i+1} - f_i}{x_{i+1} - x_i} - \frac{f_i - f_{i-1}}{x_i - x_{i-1}} + 2\lambda[\mathbf{M}^T_{ij} w_j (y_j - \mathbf{M}_{jk} f_k)] = 0 \quad i = 1,..., N \tag{3.8}$$

where the boundary conditions $f_{N+1} = f_N$ and $f_0 = f_1$ are appropriate and $\mathbf{M}^T$ is the tranpose of $\mathbf{M}$. The minimization with respect to α, β, T_1 and T_2 will be effected by a direct search for the minimum of Q in parameter space as the variational equations obtained are nonlinear.

4. Solution Procedures

The idea we have used to solve (3.8) is to notice that its structure is in some ways similar to a higher-dimensional, steady-state heat-flow problem. For the heat flow problem we know that any original distribution of temperatures will eventually decay to the steady-state distribution. One standard technique in the solution of a time independent heat-equation is to introduce a pseudo-time and integrate the pseudo-time dependent heat-flow equations forward until the steady-state temperature distribution is obtained. We use the same, pseudotime approach to solve (3.8); that is we introduce a pseudo-time which we may advance until an asymptotically steady solution is obtained which is exactly the solution of (3.8).

We seek to do this in such a way that we may factor the equation (3.8) as a product of (approximately) upper and lower triangular matrices, following the procedure of Baker and Oliphant ([1], [2], [3]). To this end let us rewrite (3.8) as

$$\begin{aligned}\gamma \mathbf{M}_{ij}^T w_j f_j - \Delta_i f + \Delta_{i-1} f - 2\lambda[\mathbf{M}_{ij}^T w_j (y_j - \mathbf{M}_{jk} f_k)] \\ + (2\lambda/\gamma)(-\Delta_i + \Delta_{i-1})\, \mathbf{M}_{ik} f_k \\ = \gamma \mathbf{M}_{ij}^T w_j f_j + (2\lambda/\gamma)(-\Delta_i + \Delta_{i-1})\, \mathbf{M}_{ik} f_k \end{aligned} \tag{4.1}$$

where we use the notation

$$\Delta_i g = \frac{g_{i+1} - g_i}{x_{i+1} - x_i} \tag{4.2}$$

We may now factor (4.1) as

$$\begin{aligned}\frac{1}{\gamma}(\gamma \mathbf{M}_{ij}^T w_j - \Delta_i + \Delta_{i-1})(\gamma \delta_{jk} + 2\lambda \mathbf{M}_{jk}) f_k^{(n)} \\ = \mathbf{M}_{ij}^T (2\lambda w_j y_j + \gamma w_j \hat{f}_j) + (2\lambda/\gamma)(-\Delta_i + \Delta_{i-1})\, \mathbf{M}_{ik} f_k^* \end{aligned} \tag{4.3}$$

where we have relabeled f_k as $f_k^{(n)}$, $\hat{f}_j$, and f_k^* in various parts of the equation. We now imagine that the first term on the left hand side and on the right side of (4.1), when taken together represent

$$\mathbf{M}_{ij} w_j \frac{\partial f_j}{\partial t} \tag{4.4}$$

where t is the pseudo-time. If we select

$$\gamma = 3/(2\Delta t), \qquad \text{and} \qquad \hat{f} = \tfrac{4}{3} f^{(n-1)} - \tfrac{1}{3} f^{(n-2)} \tag{4.5}$$

then we obtain the standard, 3-point difference approximation for the time derivative

$$\frac{\partial f^{(n)}}{\partial t} = \left(\frac{3}{2} f^{(n)} - 2f^{(n-1)} + \frac{1}{2} f^{(n-2)}\right) \Big/ \Delta t \tag{4.6}$$

for f^* we use linear extrapolation

$$f^* = 2f^{(n-1)} - f^{(n-2)}. \tag{4.7}$$

Using these approximations, the right hand side is completely known. For small Δt we have errors of only $(\Delta t)^2$. Even for large (Δt), the terms $\hat{f}$ and f^* cancel their counter-terms on the left-hand side of eq. (4.3) asymptotically as $t \to \infty$. Let us introduce

$$g_j^{(n)} = (\gamma + 2\lambda \mathbf{M}_{jk}) f_k^{(n)} \tag{4.8}$$

We obtain from (4.3) the equation for $g_j^{(n)}$ as

$$\begin{aligned}&[\mathbf{M}_{ij}^T w_j + \delta_{ij}\gamma^{-1}(\Delta_{j-1} - \Delta_j)]\, g_j^{(n)} \\ &\quad = \mathbf{M}_{ij}^T(2\lambda w_j y_j + \gamma w_j f_j) + (2\lambda/\gamma)(\Delta_{i-1} - \Delta_i)\,\mathbf{M}_{ik} f_k^*\end{aligned} \tag{4.9}$$

As $\mathbf{M}_{ij}^T$ is lower triangular, the equation for $g_j^{(n)}$ is lower triangular plus the first super diagonal. As (4.9) is linear we could make two guesses for $g_1^{(n)}$ and solve the equations $i = 1,\ldots, N-1$ for the other $g_j^{(n)}$ and then select a linear combination of these two solutions which satisfies the $i = N$ equation. Unfortunately, the magnitude of $\mathbf{M}$ relative to Δ_i can be such that an excessive amount of accuracy is lost using this procedure. We adopt an alternate procedure. We rewrite (4.9) as

$$\begin{aligned}&\left(\mathbf{M}_{ij}^T w_j + \delta_{ij}\gamma^{-1}\left(\frac{1}{x_i - x_{i-1}} + \frac{1}{x_{i+1} - x_i}\right)\right)\tilde{g}_j \\ &\quad = \gamma^{-1}\left(\frac{g_{i-1}^*}{x_i - x_{i-1}} + \frac{g_{i+1}^*}{x_{i+1} - x_i}\right) + \text{R.H.S. of (4.9)}\end{aligned} \tag{4.10}$$

where we relabeled $g^{(n)}$ as g^* and $\tilde{g}$ in part of the equation. Our solution procedure is guess g^* from f^* via (4.8), solve for $\tilde{g}$ from (4.10) as the left hand side is now triangular. Use linear extrapolation to obtain a new g^* and iterate until g^* and $\tilde{g}$ agree. The result is $g^{(n)}$. This iteration procedure (a form of relaxation) converges very rapidly for this equation, and a decrease of the error by a factor of ten or more per iteration is not uncommon. As a practical matter we required about 6 figure consistency for the solution. Having solved (4.9) by the above "minor-cycle" iteration we may now easily solve (4.8) for $f^{(n)}$ as that equation is triangular. The solution for $f^{(n)}$ from $f^{(n-1)}$ and $f^{(n-2)}$ constitutes one "major cycle" iteration. Guided by the results of Baker and Oliphant [1, 2, 3] we pick

$$\gamma = 2\lambda(1 + \alpha + \beta) + \frac{25}{N\,\text{Min}\left(w_i \dfrac{(x_{i+1} - x_i)(x_i - x_{i-1})}{x_{i+1} - x_{i-1}}\right)} \tag{4.11}$$

and again find an error reduction by a factor of 100 for each 7 to 15 "major cycles". We required an accuracy of

$$\sum_{i=1}^{N} (f_i^{(n)} - f_i^{(n-1)})^2 \leqslant 10^{-7} N \tag{4.12}$$

to terminate an iteration procedure where the f_i are normalized so that the third peak has height 10.0, in order to quote a relatively experiment-free rule.

The final variation with respect α, β, T_1 and T_2 is accomplished by use of Powell's [10] method as applied to Q.

As an illustration of our method, we have analyzed some data of Erman, Emery, and Perlman [11] on Tm^{169}. In Figure 1 we have illustrated this data together with our solution for the three component peaks. We have chosen $c(x)$ of (3.2) to be 1.0, except in the flat tail region (first 23 of 106 data points) we choose it to 100 to

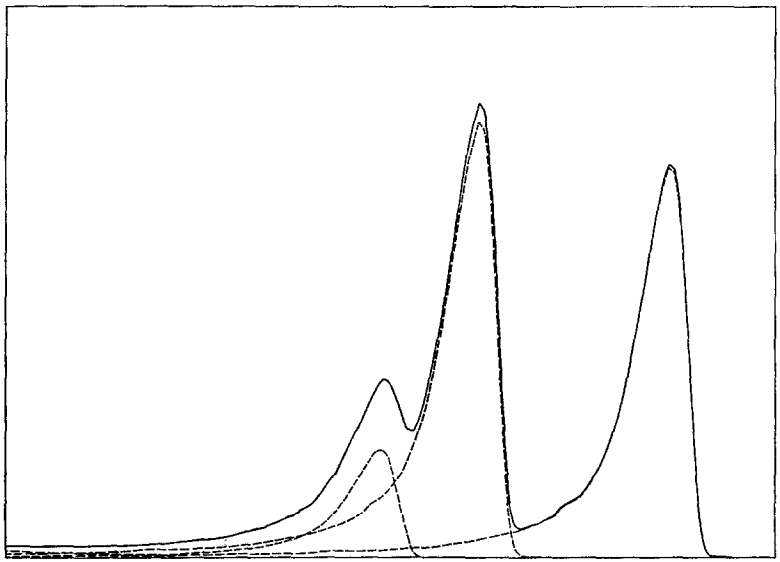

FIG. 1. A plot of the sample data (solid curve) together with the three resolved peaks (dashed lines). The right hand resolved peak and the data are virtually indistinguishable except in the tail region.

accentuate the known flatness in that region. We determine that the Lagrange multiplier λ for this case is about 0.1 in order to satisfy (3.5) with $S_0 \sim N$. We first, in order to save computer time, solved for α, β, T_1 and T_2 with $S_0 = 0(\lambda = \infty)$. It will be noted that in this limit, (3.8) can be easily cast into triangular form. We obtained $\alpha = 1.106$ and $\beta = 0.267$ as the best fit in this case. For $\lambda = 0.1$, the results change appreciably to $\alpha = 1.118$ and $\beta = 0.276$. These results are to be compared to the hand-graphical procedures previously employed [11] which yielded $\alpha = 1.14 \pm 0.02$ and $\beta = 0.292 \pm 0.005$. The use of smoothing in this problem is apparently significant and agrees well with the results of hand graphical procedures previously employed. The entire calculation for one value of λ required about 7 minutes on the CDC 6600 computer.

Very appreciable savings in computer time can be effected by recognizing that $\mathbf{M}^T$ represents shifts of the opposite sign to those in (2.1), and by using a direct

interpolation to find them, as was done to construct **M**. This procedure has the advantage that the approximate $\mathbf{M}^T$ is now comprised of narrow strips parallel to the main diagonal. It has the disadvantage that we can no longer be absolutely certain that Q is nonegative definite.

ACKNOWLEDGMENTS

We are grateful to Dr. M. Perlman whose continued interest in the solution of this problem stimulated the work reported here, and to him for making his experimental data as illustrated in Figure 1, available to us for analysis prior to publication. We are also indebted to S. Bender who programmed this procedure for the CDC 6600 computer.

REFERENCES

1. G. A. BAKER, JR. and T. A. OLIPHANT, An Implicit numerical method for solving the two-dimensional heat equation. *Quart. App. Math.* **17**, 361–373 (1960).
2. G. A. BAKER, JR., Note on the solution of the neutron diffusion problem by an implicit numerical method. *Quart. App. Math.* **17**, 314–316 (1959).
3. G. A. BAKER, JR., An Implicit numerical method for solving the n-dimensional heat equation. *Quart. App. Math.* **17**, 440–443 (1960).
4. C. DE BOOR, "Best Approximation properties of spline functions of odd degree." *J. of Math. and Mech.* **12**, 747–749 (1963).
5. B. C. COOK, Least structure analysis of photonuclear yield functions. *Nuclear Instruments and Methods.* **24**, 256–268 (1963).
6. B. ASKER, The spline curve, a smooth interpolating function used in numerical design of ship-lines. *Nordisk Tidskrit for Informations-Behandling* **2**, 76–82 (1962).
7. B. PODOLSKY and H. H. DENMAN, Conditions on minimization criteria for smoothing. *Math. Comp.* **18**, 411–447 (1964).
8. See, for example, R. L. ANDERSON and T. A. BANCROFT, "Statitistical Theory in Research," McGraw-Hill Book Company, Inc., New York, 1952.
9. See, for example, R. WEINSTOCK, "Calculus of Variations," McGraw-Hill Book Company, Inc., New York, 1952.
10. M. J. D. POWELL, An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Computer J.* **7**, 155–162 (1964).
11. P. ERMAN, G. T. EMERY, and M. L. PERLMAN, $\mathscr{L}$-subshell ratio of some low-energy E2 transitions. *Phys. Rev.*, **147**, 858–860 (1966). and private communication.