JOURNAL OF COMPUTATIONAL PHYSICS 53, 72–81 (1984)

On a Class of Least-Squares Curve-Fitting Problems

IAN J. M. BOOTH AND ANDREW D. BOOTH

Institute of Ocean Sciences, P. O. Box 6000, Sidney, British Columbia V8L 4B2, Canada

Received November 30, 1982; revised April 15, 1983

Two dielectric constant equations and the Lorentzian curve as a representation of spectral line profiles are shown to be special cases of a general curve-fitting problem. It is shown how, in these cases, the problem can be reduced to one in two variables and suggestions for appropriate computer solutions are presented. For the particular cases cited, it is shown that two are likely to lead to convergent processes, whilst the third may encounter difficulties with relative minima.

INTRODUCTION

During recent months we have been involved in the generation of least-squares fits to data from three classes of a physical problem.

First, the Lorentzian fit to a set of frequency–amplitude data whose assumed form is

$$E(\omega) = A + B/[C^2 + (\omega - \omega_0)^2] \tag{1}$$

where $E(\omega)$ is the observed amplitude at frequency ω, and A, B, C and ω_0 are constants to be so determined as to minimize

$$R = \sum_{\omega} (E(\omega) - A - B/[C^2 + (\omega - \omega_0)^2])^2.$$

The second problem occurs in dielectric constant theory where

$$E_2(\omega) = (\varepsilon_0 - \varepsilon_1)[Gf_1 + (1 - G)f_2] \tag{2}$$

and here

$$f_j = \omega T_j/[1 + \omega^2 T_j^2] \qquad (j = 1, 2)$$

and G is a constant.

The third problem is again from dielectric theory and involves the Cole–Davidson approximation [1] for complex dielectric constants:

$$E_1(\omega) - iE_2(\omega) = \varepsilon_1 + (\varepsilon_0 - \varepsilon_1)/(1 + i\omega T)^{\beta} \tag{3}$$

0021-9991/84 $3.00

where $E_1(\omega)$ and $E_2(\omega)$ are the observed quantities for angular frequency ω, and ε_0, ε_1, T and β are the constants whose optimum values are to be determined.

All of the above equations are, to some extent at least, non-linear and all involve the determination of four real constants.

The object of this note is to indicate a simple method of computation which will, generally, produce rapid convergence to the desired solution, and which requires no large main-frame computer. Rather it is well adapted to the capacity of the small "personal" computers which are now available to most scientists.

It should be noted that many other minimization procedures are available [2, 3]. The advantages of the present method are that it consumes less program storage space than those methods which involve the calculation of first, and sometimes second, derivatives. Furthermore, since it is effectively an interval halving procedure, preceded by an overall "survey," it tends not to fall into relative minima. In the latter connection, however, the last section of this paper is revealing.

Reduction to a Standard Form

We first note that Eq. (1) is of the form

$$E(\omega) = AF_1(C, D, \omega) + BF_2(C, D, \omega) \tag{4}$$

where, in this case, $F_1(C, D, \omega) = 1$. Equation (2) can be reduced to the same form by taking

$$A = (\varepsilon_0 - \varepsilon_1)G$$
$$F_1(C, D, \omega) = f_1 - f_2$$

where $C = T_1$, $D = T_2$, $B = (\varepsilon_0 - \varepsilon_1)$ and $F_2 = f_2$.

Equation (3) is somewhat more complicated. We first separate real and imaginary parts to obtain

$$E_1(\omega) = \varepsilon_1 + (\varepsilon_0 - \varepsilon_1)(1 + \omega^2 T^2)^{-\beta/2} \cos(\beta\phi)$$
$$E_2(\omega) = (\varepsilon_0 - \varepsilon_1)(1 + \omega^2 T^2)^{-\beta/2} \sin(\beta\phi)$$

where

$$(1 + \omega^2 T^2)^{1/2} \exp(i\phi) = 1 + i\omega T \tag{5}$$
$$\tan\phi = \omega T \tag{6}$$

whence, using (5) again,

$$\begin{aligned} E_1(\omega) &= \varepsilon_1 + (\varepsilon_0 - \varepsilon_1)(\cos\phi)^\beta \cos(\beta\phi) = \varepsilon_1 + (\varepsilon_0 - \varepsilon_1)\, C_1(\omega) \\ E_2(\omega) &= (\varepsilon_0 - \varepsilon_1)(\cos\phi)^\beta \sin(\beta\phi) = (\varepsilon_0 - \varepsilon_1)\, C_2(\omega) \qquad \text{(say).} \end{aligned} \tag{7}$$

We see that, to effect a least-squares fit, we must minimize

$$R = \sum_{\omega} [E_{10}(\omega) - \varepsilon_1 - (\varepsilon_0 - \varepsilon_1) C_1(\omega)]^2 + [E_{20}(\omega) - (\varepsilon_0 - \varepsilon_1) C_2(\omega)]^2$$

where $E_{10}(\omega)$ and $E_{20}(\omega)$ denote the experimentally determined values.

The values of ε_0 and ε_1 are obtained from

$$\frac{\partial R}{\partial \varepsilon_0} = -2 \sum_{\omega} C_1(\omega)[E_{10}(\omega) - \varepsilon_1 - (\varepsilon_0 - \varepsilon_1) C_1(\omega)] + C_2(\omega)[E_{20}(\omega) - (\varepsilon_0 - \varepsilon_1) C_2(\omega)] = 0$$

and

$$\frac{\partial R}{\partial \varepsilon_1} = -2 \sum_{\omega} (1 - C_1(\omega))[E_{10}(\omega) - \varepsilon_1 - (\varepsilon_0 - \varepsilon_1) C_1(\omega)] - C_2(\omega)[E_{20}(\omega) - (\varepsilon_0 - \varepsilon_1) C_2(\omega)] = 0.$$

It is clear that these two equations are linear in ε_0 and ε_1 and can easily be solved explicitly to give their optimum values in terms of ω and β if these were known.

An exactly similar argument applies to our "standard" form (4). Here

$$R = \sum_{\omega} [E(\omega) - AF_1 - BF_2]^2$$

whence A and B are determined from

$$\frac{\partial R}{\partial A} = -2 \sum_{\omega} F_1 [E(\omega) - AF_1 - BF_2] = 0$$

$$\frac{\partial R}{\partial B} = -2 \sum_{\omega} F_2 [E(\omega) - AF_1 - BF_2] = 0$$

which lead to

$$\begin{aligned} A \sum_{\omega} F_1^2 + B \sum_{\omega} F_1 F_2 &= \sum_{\omega} F_1 E(\omega) \\ A \sum_{\omega} F_1 F_2 + B \sum_{\omega} F_2^2 &= \sum_{\omega} F_2 E(\omega) \end{aligned} \tag{8}$$

and thus to

$$\begin{aligned} A &= \frac{\sum_\omega F_2^2 \cdot \sum_\omega F_1 E(\omega) - \sum_\omega F_1 F_2 \cdot \sum_\omega F_2 E(\omega)}{\sum_\omega F_1^2 \cdot \sum_\omega F_2^2 - (\sum_\omega F_1 F_2)^2} \\ B &= \frac{\sum_\omega F_1^2 \cdot \sum_\omega F_2 E(\omega) - \sum_\omega F_1 F_2 \cdot \sum_\omega F_1 E(\omega)}{\sum_\omega F_1^2 \cdot \sum_\omega F_2^2 - (\sum_\omega F_1 F_2)^2}. \end{aligned} \tag{9}$$

A Useful Approximation for the Cole–Davidson Case

Inspection of typical plots of $E_2(\omega)$ vs ω shows the general form of Fig. 1.

It is of interest to investigate the location of the maximum. At this point we have $dE_2/d\omega = 0$. Now

$$\frac{dE_2}{d\omega} = \frac{dE_2}{d\phi} \cdot \frac{d\phi}{d\omega}$$

and, using (6) and (7), we have

$$\frac{d\phi}{d\omega} = T \cos^2 \phi$$

and now

$$\frac{dE_2}{d\phi} = (\varepsilon_0 - \varepsilon_1)(-\beta \cos^{\beta-1}\phi)(\sin \phi \sin \beta\phi - \cos \phi \cos \beta\phi)$$

whence

$$\frac{dE_2}{d\omega} = T\beta(\varepsilon_0 - \varepsilon_1) \cos^{\beta+1}\phi \cos|(\beta + 1)\phi|.$$

Thus, for $dE_2/d\omega = 0$ one of the following must be true:

(i) $\cos \phi = 0$,

(ii) $\beta = 0$,

(iii) $\cos((\beta + 1)\phi) = 0$.

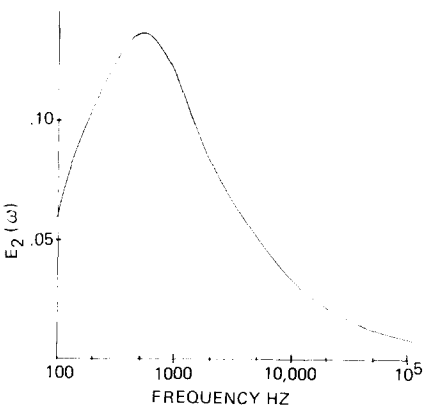

FIG. 1. A typical plot of $E_2(\omega)$ vs frequency showing the existence of a maximum.

(i) leads to $\omega T = \infty$, which is clearly unacceptable, (ii) is similarly unrewarding, but (iii) gives

$$(\beta + 1)\tan^{-1}(\omega T) = \pi/2$$

or

$$T = \tan(\pi/2(\beta + 1))/\omega. \tag{10}$$

Since, in practice, β varies only slowly with frequency for a given substance, Eq. (8) with a mean value of β allows an adequate initial estimate of T to be made. In some data supplied to us by Dr. S. Walker [4] it transpired that $\beta = 0.65$ provided a useful starting value.

The Minimization Process

On a large and fast machine it might prove profitable to apply some direct iterative procedure to the solution of the equations

$$\frac{\partial R}{\partial A} = 0, \qquad \frac{\partial R}{\partial B} = 0, \qquad \frac{\partial R}{\partial C} = 0, \qquad \frac{\partial R}{\partial D} = 0$$

for example, the method of steepest descents [5, 6]. However, even for relatively simple forms of $F_1(C, D, \omega)$, $F_2(C, D, \omega)$ the algebra becomes formidable and, with the interpreter types of operating system used in small computers, the running time is usually prohibitive.

Instead we have found that a simple procedure, based upon R itself, is fast and adequate. Consider first a single dimension, and assume that an approximation x_0 to the maximum of $R(x)$ is known. It is easily shown [7] that if $R(x)$ is evaluated at x_{-1}, x_0 and x_{+1}, where $x_0 - x_{-1} = x_1 - x_0 = h$ (say), then a better approximation to the minimum, or maximum, is

$$x = x_0 + \frac{R(x_{-1}) - R(x_1)}{2[R(x_{-1}) - 2R(x_0) + R(x_{+1})]} \cdot h. \tag{11}$$

In two dimensions the situation is slightly more complicated. Several strategies are available but the one which gave the best performance in terms of minimizing machine operations (and hence running time) and, at the same time, having an acceptable rate of convergence is derived from a well-known two-dimensional interpolation formula [8], and it is based on the sampling scheme shown in Fig. 2.

If we consider a function $R(x, y)$ to have a minimum in proximity to (x_0, y_0), then, defining

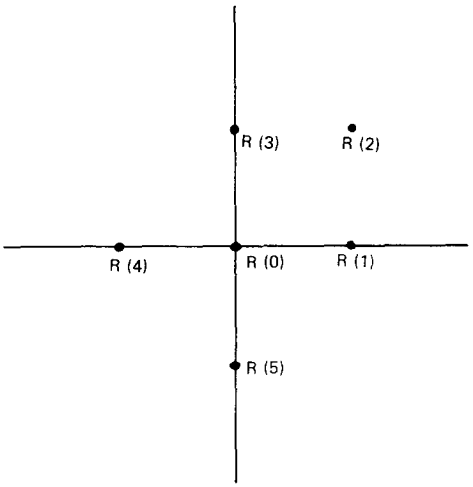

FIG. 2. Sampling scheme for residuals.

$$\delta^2 x = R(1) - 2R(0) + R(4)$$
$$\delta^2 y = R(3) - 2R(0) + R(5)$$
$$\delta xy = R(0) - R(1) + R(2) - R(3)$$
$$\Delta = 2[\delta^2 x\, \delta^2 y - (\delta xy)^2]$$

a better approximation to the minimum (measured from $R(0)$) has coordinates

$$x = \frac{[R(4) - R(1)]\, \delta^2 y - [R(5) - R(3)]\, \delta xy}{\Delta} \cdot h$$
$$y = \frac{[R(5) - R(3)]\, \delta^2 x - [R(4) - R(1)]\, \delta xy}{\Delta} \cdot k \qquad (12)$$

where h and k are the respective x and y steps.

The iterative solution process is thus:

1. Establish initial estimates of the parameters, by inspection or otherwise.

2. Using Eqs. (9) together with the equation for R, calculate the values of the residual at the points $R(0) \cdots R(5)$ defined in Fig. 2.

3. Using Eqs. (12) calculate new approximations to C and D.

4. Recalculate $R(0)$ at the new trial point. If $R(0)$ is acceptably small then end. Or else, repeat the process from step 2 reducing step lengths if required.

An Actual Programme

To show the simplicity of the process, we present, in Table I, an actual programme, written in BASIC, for investigating the convergence of the two time constant case of Eq. (2). Since this programme was designed for convergence investigations the actual function values are calculated (lines 80–90) rather than observed as would normally be the case.

The programme is designed so as not to require initial parameter investigation. This is achieved by a "general survey" over all realistic parameter values over a grid (lines 100–190). Since the frequency scale of real dielectric observations is exponential, to effect a suitable compression, logarithms are taken at line 110 and reconverted at line 130. A subroutine is used (lines 440–510) to evaluate the residue R and the survey programme simply takes as the best starting values those producing the smallest R in the field.

In connection with this subroutine it should be noted that a simple algebraic manipulation enables the residual to be calculated in a single pass through the data. That is, it is not necessary to first calculate values of A and B as described by Eq. (9) and then compute R in a second iteration. It is easily shown that

$$R = \sum_{\omega} E(\omega)^2 - A \sum_{\omega} F_1 E(\omega) - B \sum_{\omega} F_2 E(\omega) \tag{13}$$

and notice that $\sum_{\omega} E(\omega)^2$ (=EE in the programme) needs to be calculated only once and this is done in line 90 of the programme. The advantage of using Eq. (13) is that it enables R to be calculated in a single iterative cycle, rather than in the two cycles which are needed if A and B are calculated first and then a direct calculation of R is made *ab initio*. This maximizes the speed of execution. The disadvantage lies in the fact that, using (13), R may result from the subtraction of large quantities and this may lead to loss of significance in the result when R is itself small. We have not encountered this problem and Eq. (13) is used in line 500 of the subroutine.

The alternative form, which does not use Eq. (13), is listed as lines 5000–5010. When the programme is run on a PET 2001 computer, the running time, for the parameters given in the programme, is 4 minutes 10 seconds. If the alternative lines are substituted for lines 500–510 the running time is increased to 6 minutes 35 seconds.

The approximations having been obtained, the refinement process described above is initiated at line 200. Note the use of Eq. (10) at lines 240–330, and the step length reduction at line 350.

The remainder of the programme is concerned with merely stoping at the preassigned accuracy using the convergence criterion $H2 <= 1E-9$ on line 370, and then to print the parameters (lines 390–400) and finally a table of calculated values (lines 410–430).

Similar programmes have been written for the Lorentzian and complex dielectric constant case.

TABLE I

```
 10 PRINT" DUAL RELAXATION TIME DIELEC.L.S. FIT"
 20 PRINT"  ENTER TRIAL PARAMETERS IN LINE 50
 30 PRINT"ENTER FREQUENCY DATA IN STATEMENTS 540-"
 40 PRINT"   PRINTS MAY BE MACHINE DEPENDANT"
 50 G=.25:T1=9.4408172E-5:T2=8.39888737E-5:PI=3.1415326 :N=10
 60 PRINT"TO STOP AT ANY STAGE PRESS SPACE & WAIT@"
 70 DIMEE2(N),F(N),X1(N),X2(N)
 80 FORI=1TON:READF:F(I)=2*π*F:W1=F(I)*T1:W1=W1/(1+W1*W1):W2=F(I)*T2
 90 W2=W2/(1+W2*W2):E2(I)=G*W1+(1-G)*W2:EE=EE+E2(I)*E2(I):NEXT
100 PRINT"MAKING GENERAL SURVEY BETWEEN 1E-6 &1E-3@
110 L=1E-6:U=1E-3:RA=3.5:R0=1000:L=LOG(L):U=LOG(U):ME=(L+U)/2
120 L1=ME-RA:U1=ME+RA:L2=L1:U2=U1:RA=RA/2
130 FORI1=L1TOU1STEPRA:D2=EXP(I1):FORJ1=L2TOU2STEPRA
140 D3=EXP(J1):GOSUB440:IFR<R0THENR0=R:DL=D2:DU=D3
150 NEXT:NEXT:GETA$:IFA$<>""THEN130
160 PRINT"T1=";DL:PRINT"T2=";DU:PRINT"R=";R0:PRINT
170 RA=RA/2:IFRA<1E-5 THEN190
180 L1=LOG(DL)-RA:U1=L1+2*RA:L2=LOG(DU)-RA:U2=L2+2*RA:GOTO130
190 C2=DL:C3=DU
200 PRINT"STARTING REFINEMENT PROCESS@
210 H2=C2-1E-7:H3=C3-1E-7:L=0
220 D2=C2:D3=C3
230 GOSUB440:R0=R
240 PRINT"R0=";R0;"H=";H2:D2=C2-H2:GOSUB440:R4=R
250 D2=C2+H2:GOSUB440:R1=R
260 DX=R4-2*R0+R1
270 D3=C3+H3:GOSUB440:R2=R
280 D2=C2:GOSUB440:R3=R
290 D3=C3-H3:GOSUB440:R5=R
300 DY=R3-2*R0+R5
310 DL=R0-R1-R3+R2
320 DN=2*(DX*DY-DL*DL)
330 PX=((R4-R1)*DY-(R5-R3)*DL)/DN:QY=((R5-R3)*DX-(R4-R1)*DL)/DN
340 D2=C2+H2*PX:D3=C3+H3*QY:GOSUB440
350 IFR>=R0THENH2=H2/8:H3=H3/8:D2=C2:D3=C3:GOTO370
360 C2=D2:C3=D3:R0=R
370 IFH2>1E-9THENL=L+1:PRINT"L=";L:GETA$:IFA$=""THENH1=H*1E-1:GOTO240
380 G=A1/A2:PRINT
390 PRINT"E(0)-E(INF)=";A2:PRINT"G=";G:PRINT"T1= ";D2:PRINT"T2=";D3
400 PRINT"H=";H2:PRINT
410 PRINT:PRINT" FREQ      E''":FORI=1TON:F=F(I)/(2*PI):W1=F(I)*D2:W2=F(I)*D3
420 W1=W1/(1+W1*W1):W2=W2/(1+W2*W2):E2=A1*W1+(A2-A1)*W2
430 PRINTINT(F),E2:NEXT:END
440 P1=0:P2=0:F1=0:F2=0:Q=0:R=0:IFD3=D2THEND3=D2+1E-7
450 FORJ=1TON:W1=F(J)*D2:W2=F(J)*D3:W1=W1/(1+W1*W1):W2=W2/(1+W2*W2)
460 W1=W1-W2
470 P1=P1+E2(J)*W1:P2=P2+E2(J)*W2:F1=F1+W1*W1:F2=F2+W2*W2:Q=Q+W1*W2
480 NEXT:A3=F1*F2-Q*Q
490 A1=(P1*F2-P2*Q)/A3:A2=(P2*F1-P1*Q)/A3
500 R=EE-A1*P1-A2*P2
510 RETURN
520 DATA1E2,2E2,5E2,1E3,2E3,5E3,1E4,2E4,5E4,1E5
530 :
540 :
600 REM BELOW ARE ALTERNATEIVE LINES   500-510 FOR THE DIRECT EVALUATION OF R
610 :
620 REM TO USE, DELETE TERMINAL  ZERO FROM THE LINE NUMBER AND RE-RECORD
630 :
640 :
5000 R=0:FORJ=1TON:W1=F(J)*D2:W2=F(J)*D3:W1=W1/(1+W1*W1):W2=W2/(1+W2*W2)
5100 W1=W1-W2:E=E2(J)-A1*W1-A2*W2:R=R+E*E:NEXT:RETURN
READY.
```

Observations on the Use of the Method

It has been found, in extensive tests, that the process works well for each of the physical problems described at the start of this paper, and converges to a minimum in fewer than 20 iterations. The programme has been used on a CBM-PET 2001, a VIC 20 (with 8K expansion), an Apple II2 and a TRS-80/III computer, and the running times were all less than 5 minutes.

An alternative programme, written in FORTRAN 80 for a TRS-80, Model III, had running times of less than 2 minutes.

Whilst convergence was always observed, however, it was found that when T_1 approaches 80% of T_2 in Eq. (2), the process converges to "incorrect" values.

On further investigation it was found that the "incorrect" value actually gave a zero value for the residual, to the 9D accuracy of the calculations. This prompted a further examination of the nature of the field of residuals in all three cases. To do this, three-dimensional plots of the residual fields were made using a large computer. The results, for the Lorentzian and for the complex dielectric cases, show the presence of well-defined, and unique, minima.

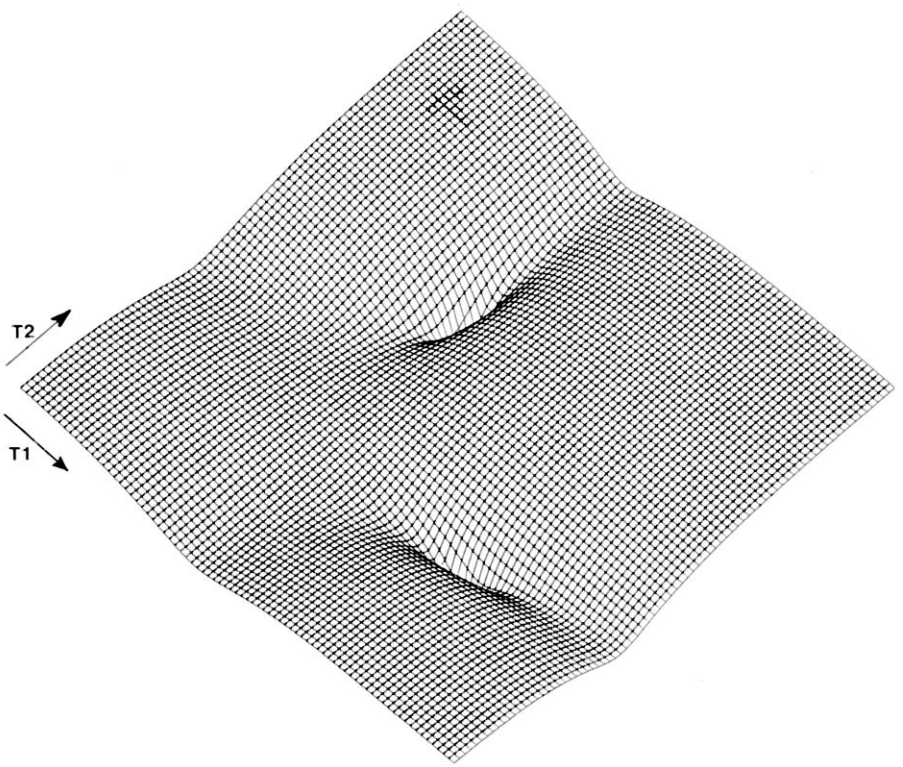

FIG. 3. Three-dimensional plot of a typical residual field for the two relaxation time problem. Note the slit-like depressions at the minima.

When the situation modelled by Eq. (2) is investigated, however, a different situation prevails. This is shown in Fig. 3. First, it is evident from Eq. (2) that there is symmetry about $T_1 = T_2$, and this guarantees the presence of at least two minima, but worse is to come. The region of the minimum is seen to be a slit-like depression and closer examination reveals the presence of relative minima along its length when T_1 approaches T_2 in value. It follows that no method of fitting based on minimization can be guaranteed to produce correct results as T_1 approaches T_2.

In adapting the programme of Table I for use with observational data the following points may be of interest.

First, it is useful to store the data in a disc (or tape) file. Not only does this provide a permanent record, but it also enables an adequate editing routine to be used. Second, in the case of the Lorentzian and Cole–Davidson formulae, a general survey is not necessary. In place of this Eq. (11) can be used, in conjunction with the data in the region of the maximum, to estimate the initial trial for ω. In the Lorentzian case this is used directly, and for the Cole–Davidson formula Eq. (10) then gives a trial value of T.

Finally, if readers are interested in having a printout of a fully operational version of the Cole–Davidson programme suitable for a CBM-PET-VIC series machine they should write to the authors.

ACKNOWLEDGMENTS

The authors wish to thank Dr. S. Walker and Mr. R. Gourlay of the Department of Chemistry, Lakehead University, Thunder Bay, Ontario, for bringing the interesting dielectric problems to their notice. They would also like to thank the referees for numerous clarifications that have made the paper much more readable.

REFERENCES

1. K. S. COLE AND R. H. COLE, *J. Chem. Phys.* **9** (1941), 341.
2. B. CARNAHAN AND J. O. WILKES, "Digital Computing and Numerical Methods," pp. 405–433, Wiley, New York, 1973.
3. G. PECKHAM, *Comput. J.* **13** (1970), 418–420.
4. S. WALKER, private communication.
5. A. D. BOOTH, *Quart. J. Mech. Appl. Math.* **2** (1949), 460–468.
6. A. D. BOOTH, *Proc. Roy. Soc. London* **197** (1949), 336–355.
7. A. D. BOOTH, "Fourier Technique," pp. 62–65, Cambridge Univ. Press, London/New York, 1948.
8. M. ABRAMOWITZ AND I.A. STEGUN, "Handbook of Mathematical Functions," N.B.S. Appl. Math. Series No. 55, 882 item 25.2.67, Washington, D. C., 1970.