

THE APPLICABILITY AND POWER OF COMPLEX NONLINEAR LEAST SQUARES FOR THE ANALYSIS OF IMPEDANCE AND ADMITTANCE DATA

J. ROSS MACDONALD

Department of Physics and Astronomy, University of North Carolina, Chapel Hill, NC 27514 (U.S.A.)

J. SCHOONMAN

Solid State Department, Physics Laboratory, State University of Utrecht, 3508 TA Utrecht (The Netherlands)

A.P. LEHNEN

Department of Physics and Astronomy, University of North Carolina, Chapel Hill, NC 27514 (U.S.A.)

(Received 7th April 1981; in revised form 8th September 1981)

ABSTRACT

An analysis is made of the ability of nonlinear complex least squares fitting to yield useful estimates of parameter values occurring in equivalent circuits. In addition, three-dimensional perspective plots are presented which show the full frequency response of either impedance or admittance data. Three different kinds of data are studied using these two methods.

For a given set of four circuit elements (two resistors and two capacitors) artificially computer-generated impedance and admittance values are constructed for a Voigt circuit having two time constants. These exact results are then contaminated either by digit round-off or addition of zero mean, normally distributed random noise. A complex least squares analysis is then made on the contaminated data and the resulting estimates of the circuit parameters and their associated standard errors are compared to the exact starting values. For the case when one of the capacitors is vanishing small, the optimal range of measurement over the resulting semi-circle in the complex impedance plane is discussed in detail. For this case the accuracy of input data required to resolve the values of the two resistors when they are separated in size by two orders of magnitude is determined. The other extreme of two time constants which are not well separated is also studied. Again one obtains approximate estimates of both the range of frequencies and the inherent accuracy of the data necessary to adequately resolve all four circuit parameters.

The actual frequency dependent admittance of a real ladder network of resistors and capacitors having three different time constants was measured on a Solartron type 1172 response analyzer. These data were then analyzed using the complex least squares procedure and the fitted circuit parameters were all in close agreement with their nominal experimental values.

Finally, the admittance response of $\beta\text{-PbF}_2$ at 474 K was measured on the Solartron 1172. A complex least squares analysis of these data yields a good fit when an equivalent circuit employing a constant phase element is used. A three-dimensional perspective plot showing very clearly the agreement between fitted and measured values of the impedance is shown.

I. INTRODUCTION

The estimation of parameter values entering theoretical models or equivalent circuits from impedance and/or admittance data is important in obtaining quantitatively accurate representations of the frequency response of liquid or solid electrolytes. Such analyses have long been used [1] and occur throughout the literature [2]. However, the methods employed in actually estimating the parameters and, in particular, the associated error distribution of the parameters, have often been quite crude. These latter error estimates are often more significant than even the parameter values themselves in determining which theoretical model is most appropriate for a given set of data.

The aim of this paper is to investigate the applicability, power, and types of results that a nonlinear complex least squares minimization [3,4] can give for several kinds of data. Here the word complex is taken to mean that both real and imaginary components of the data are simultaneously fitted in the least squares minimization [4-6]. In addition, we demonstrate the value of three-dimensional perspective plotting of impedance and/or admittance data as opposed to ordinary complex plane plotting.

Several recent studies [7-10] have been made which involve relatively simple, non-least squares determinations of model parameters. All of these, however, have used electrochemical data taken on actual, distributed systems. Thus, a fundamental understanding of the inherent ability of the method of analysis to resolve different parameter values may remain obscured or exaggerated by the presence of systematic errors which inevitably appear in fitting measurements on such real electrolytic systems. To clarify this aspect, in Section II several least squares analyses are performed on artificial, computer-generated circuit data which have been contaminated by either digit round-off or the addition of zero mean, normally distributed random error to each data value. The circuits analyzed consist of simple Voigt networks of ideal resistors and capacitors. In Section III, complex least squares fits of actual frequency response data for two different experimental systems are carried out and discussed. The first such system measured involved an actual ladder network of resistors and capacitors having three different time constants, while the second one was a sample of β -PbF₂ at 474 K. In this latter case, a three-dimensional plot (Real Z , $-\text{Im}Z$, $\log f$) showing the agreement between fitted and measured values of the impedance Z as a function of the frequency f is also presented.

II. SYNTHETIC DATA

In general the frequency response of a system does not uniquely determine the circuit used to model that response. For example, in Fig. 1, three different circuits are shown. Each circuit has six lumped elements (three ideal resistors and three ideal capacitors) and three ($N=3$) time constants. The total impedance for all three circuits will be a polynomial in angular frequency, ω , of degree four divided by a polynomial of degree three. For any given choice of the six elements for one of the

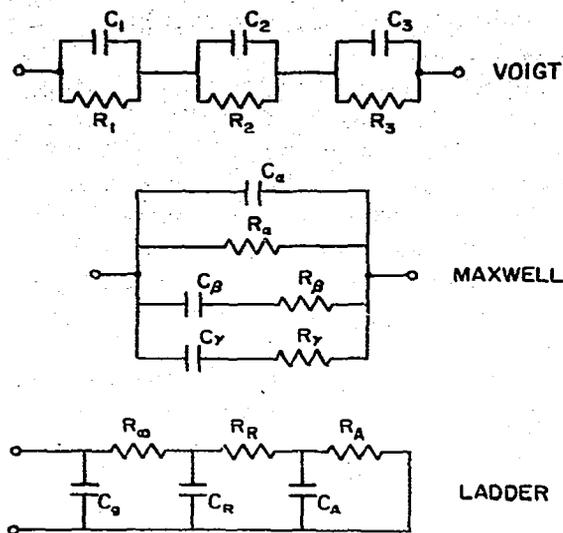


Fig. 1. Three $N=3$ circuits which can have the same impedance-frequency relation.

models, there exist corresponding sets of circuit elements for each of the two remaining networks such that all three have exactly the same impedance at all frequencies. The choice here of circuits having three time constants is unimportant, and such an equality can be achieved for any N [11,12]. Starting from the Voigt network, it is possible to obtain simple algebraic formulae for the corresponding Maxwell or ladder circuit parameters [13].

However, it seems reasonable for any specific real extended physical system that the actual response is best represented by a unique circuit. Different theoretical response models can lead to different equivalent circuits, such as any of the three of Fig. 1. Even when one believes that a particular model and circuit is appropriate, one may be wrong, and it is thus desirable to have means available to test the question. One way to discover which of several circuits is most appropriate is to repeat measurements over a variety of different conditions (e.g., temperature or electrode separation). For a less appropriate circuit, most if not all of the parameter values will then vary appreciably, while for the most appropriate model some of the parameters may not vary at all, and most of the variation would be expected to occur in a minimum set of remaining parameters. For example, if all but a single parameter of the ladder network are taken to be independent of temperature or electrode separation, fitting of impedance data derived from this circuit, for say two different separations, using either of the other circuits of Fig. 1 will yield equivalent overall fits but fits in which, in general, not just one but all or most of the parameters values estimated from the fitting will be found to be separation dependent. Thus, a single fitting of a set of data to a given equivalent circuit will not generally yield estimates of the most appropriate parameters unless the equivalent

circuit has already been proven to be the most appropriate of those with the same impedance-frequency relation for the situation under study.

Artificial computer generated impedance data was constructed for a two-time-constant Voigt network having $(R_1, C_1; R_2, C_2) = (10 \Omega, 0; 10^3 \Omega, 1 \mu\text{F})$. The resulting three-dimensional (conjugate) impedance (Z^*) plot is given in Fig. 2. The origin of the $\log(f)$ scale is taken at -2 for these data. The real and imaginary parts of the impedance are plotted in units of 200Ω . The conventional two-dimensional Z^* plot (here the back projection) is just a semi-circle centered on the real axis at $Z^* = R_1 + (R_2/2)$ and having radius $(R_2/2)$. The capacitance C_2 manifests itself in this figure only in the frequency dependence shown along the $\log(f)$ axis. For these data 145 points were generated starting at a frequency of 1.389×10^{-2} Hz and ending at 1.824 MHz. For the conventional two-dimensional Z^* plot, this means that the first point on the semi-circle starts at an angle of 0.01° and the points extend to an angle of 179.99° . Coverage of the 145 points was such that they are more or less equally spaced in angle with the angular separation of adjacent points being about 1.25° .

Following Despic et al. [8], for the purposes of data analysis one can consider the semi-circle being divided into three sections: the first, Sector 1, from 0° to 30° , the second, Sector 2, from 61.25° to 120° , and the last, Sector 3, from 151.25° to 180° . For the data shown in Fig. 2, Sector 1 contains 25 points, Sector 2 48 points, and Sector 3 24 points. An additional sector covering the whole semi-circle, but skipping every other value (thus, having a spacing of about 2.5°) was designated as Sector 4 and contained 73 points. In Fig. 2, Sector 1 starts with the first point and goes up to the first double projection into the $\log(f)$, $\text{Re}(Z)$ plane. Sector 2 starts at the second double projection and goes up to the third. All the points after the last double projection constitute Sector 3.

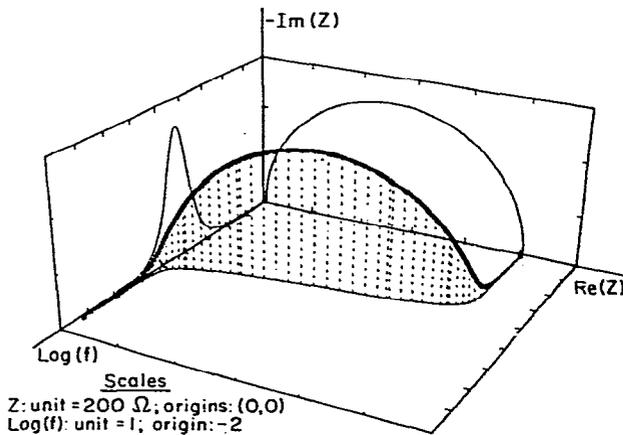


Fig. 2. Perspective three-dimensional plot of the Z^* response of a $N=2$ Voigt circuit when $(R_1, C_1; R_2, C_2) = (10 \Omega, 0; 10^3 \Omega, 1 \mu\text{F})$. Individual data points are represented by small squares and connected by a continuous curve, while planar projections are shown only as continuous curves.

To simulate experimental error in these exact (actually 13-digit) data, the number of digits was rounded to n ($n \leq 4$). Then, a complex least squares determination of R_1 , R_2 , and C_2 was made for a variety of sectors and choices of n . The virtues and advantages of a least squares analysis are well known [4,14]. Among these are that the resulting solution, even for a non-Gaussian distribution of error, is both unbiased and has minimum variance. For the case of impedance and/or admittance data a complex least squares fit puts both real and imaginary components on an equal footing (apart from any inequities brought about by weighting differences). The results of the analysis are estimated parameter values and their estimated uncertainties, along with an estimate of the overall standard deviation of the fit, σ_f . Thus, if the data determine some parameter values better than others, this is immediately obvious. Furthermore, in the analysis of real extended systems equivalent circuit types such as those of Fig. 1 should always lead to the same σ_f . However, the relative parameter uncertainties will in general differ, and they thus will help provide information for discriminating between different models.

Let the complex function, $f(\omega; \bar{R}) \equiv f_x(\omega; \bar{R}) + if_y(\omega; \bar{R})$, represent the impedance (Z) or admittance (Y) response of the model or equivalent circuit used to describe the system. Here f is a function both of the angular frequency, ω , and the set of model parameters, \bar{R} . For the purposes of this fit, the frequency, as is usually the case experimentally, is assumed to have negligible error as compared to the "measured" values of Z (or Y). In this approximation, with measurements made at angular frequencies ω_i ($i = 1, \dots, K$) of complex Z (or Y) data of the form $x_i + iy_i$, the goal of a complex least squares analysis is to find that set of parameters which minimizes the weighted sum

$$S = \sum_{i=1}^K \left\{ w_i^x [x_i - f_x(\omega_i; \bar{R})]^2 + w_i^y [y_i - f_y(\omega_i; \bar{R})]^2 \right\} \quad (1)$$

Here w_i^x and w_i^y are the weights associated with the i th data point. If the standard deviations σ_i^x and σ_i^y of each point are known, then $w_i^x = (\sigma_i^x)^{-2}$ and $w_i^y = (\sigma_i^y)^{-2}$. The choice of $w_i^x = w_i^y = w$, where w is a positive number independent of i , is equivalent (as far as the choice of R for which S is a minimum) to unity weighting (i.e., unweighted data).

The initial least squares determination of parameters was done by a computer program employing the Levenberg-Marquardt [15,16] algorithm. This part of our code was supplied by the Argonne Laboratory and was written by B.S. Garbow, K.E. Hillstrom and J.J. Moré. The final output value of \bar{R} (the converged parameter set) of this part of the total program was then used as the starting input parameters for a separate computer code which calculated, among other things, the parameter covariance matrix. This provided us with a statistical estimate of the error distribution of the parameters. This program is based on a nonlinear least squares package given to us by J.D. Olson of Union Carbide Corporation and written by H.I. Britt and R.H. Luccke [6]. All of these calculations were done on an IBM 360/370 operating system at the University of North Carolina Computation Center.

The complex least squares analyses were carried out on both unweighted (or unity weighted) data, designated by $W \equiv 0$, and, in addition, with the assumption that the random errors in the data were proportional to the magnitudes of the data values ($W \equiv P$). The results are given in Table 1.

Here the numbers following the \pm signs are the least squares estimated parameter uncertainties. As one would expect from Fig. 2, the ability to resolve R_1 (here $R_1/R_2 = 10^{-2}$) is most difficult if the data only samples the low frequency tail of Sector 1. This is true even for $n = 3$ (i.e., data accurate to about 0.1%). If data from Sectors 2 or 3 are used, even the $n = 2$ case gives parameters which are within about a percent (the "inherent" accuracy of the data) of the "actual" values. It is also interesting to note that Sector 2, again as one might have predicted from Fig. 2, is the optimal sector. Since it samples the central portion of the semicircle, it gives good estimates of all three parameters, even for $n = 2$. In fact, as Table 1 shows, the $n = 2$ data analyzed over Sector 2 give slightly better results than Sector 1 data which is accurate to 0.01% ($n = 4$)! Perhaps even more surprising is that the analysis for $n = 2$ on Sector 2 appears to give results at least as good as the analysis made over the whole frequency range (Sector 4). Presumably, this could be interpreted as meaning that with an accuracy of only about 1% ($n = 2$) one should in fact avoid impedance data from either end of the semicircle, where errors are disproportionately exaggerated. Finally, one sees that the parameter uncertainty estimates for Sections 1 and 3 are often too small. In contrast, the analysis of Sector 2 data gives results which are within 2 estimated parameter uncertainties of the "correct" values. The general effect of the $W = P$ vs. the $W = 0$ weighting is to improve both parameter and parameter uncertainty estimates.

The fact that data from the middle portion of the impedance semicircle is the most reliable in estimating parameter values has also been noted by Despic and co-workers [8]. However, because their treatment was not a full least squares determination and, in particular, since their estimation of parameter uncertainties

TABLE 1

Complex least squares fitting results for the synthetic impedance data of Fig. 2 when error is introduced by digit roundoff

W	Sector	n	R_1/Ω	R_2/Ω	$C_2/\mu F$
0	1	4	10.56 \pm 0.11	999.45 \pm 1.03	1.0011 \pm 0.0023
0	1	3	22.23 \pm 9.95	988.17 \pm 9.70	1.0245 \pm 0.0206
P	1	3	15.80 \pm 6.04	994.43 \pm 5.88	1.0110 \pm 0.0122
0	2	3	10.26 \pm 0.14	999.95 \pm 0.08	1.0004 \pm 0.0003
0	3	3	10.01 \pm 0.05	999.33 \pm 1.70	1.0004 \pm 0.0003
0	4	3	10.01 \pm 0.10	1000.24 \pm 0.12	0.9999 \pm 0.0003
P	1	2	46.54 \pm 20.1	961.87 \pm 19.6	1.0828 \pm 0.0451
0	2	2	10.14 \pm 1.24	1000.39 \pm 0.70	1.0003 \pm 0.0026
0	3	2	9.84 \pm 0.16	985.22 \pm 5.20	1.0024 \pm 0.0008
0	4	2	9.65 \pm 0.49	998.32 \pm 0.57	0.9982 \pm 0.0014
P	4	2	9.89 \pm 0.04	999.21 \pm 0.80	0.9999 \pm 0.0009

was non-statistical, they seem to have underestimated the region of parameter space where a good resolution of all parameter values can be obtained. This explains why our Sector 2 analysis of data with n as small as 2 would seem to contradict their conclusion that acceptable accuracy in the determination of R_1 is obtainable only when R_1 is not more than one order of magnitude smaller than R_2 .

It is interesting to note that Despici et al. claim to have done a statistical determination of R_1 and R_2 by a least squares fit of the impedance data to a semicircle (*not* complex least squares). We too have considered this problem, and show in Appendix A how a generalized inverse technique can give excellent starting values for both the radius and the circle center.

Since the minimization of eqn. (1) can be done for either Z or Y data, it is of interest to see how this choice of measurement type affects the parameter estimates. Accordingly, the exact (thirteen digit) impedance data of Fig. 2 was inverted to yield Y results. The grouping of points into the various sectors was left unchanged. Errors in the data were again simulated by rounding to n digits. The results of a complex least squares determination of R_1 , R_2 , and C_2 for various values of n and different sectors are presented in Table 2. Here the $W=P$ and $n=2$ analysis of Sector 1, which was presented in Table 1, is not given. The reason for this omission is that although the Levenberg-Marquardt algorithm converged to a parameter set having $(R_1; R_2, C_2) = (0.21 \Omega; 1009.89 \Omega, 0.9821 \mu\text{F})$, the Luecke-Britt routine did not converge. Thus, parameter uncertainty estimates on all three parameters were not obtained. From previous fitting work this seems to illustrate that, while the Levenberg-Marquardt algorithm seems more robust (i.e., it will converge for a very large class of problems) and is very fast in terms of computing time, it sometimes converges to local minimum solutions which are not fully reliable (not the absolute minimum of the sum of squares). Thus, our use of the second independent analysis on the results of the first gives us a powerful check on the reliability and stability of the estimated parameter values.

From Table 2 one sees that the use of Y data, like the Z fitting, gives the worst results when only Sector 1 is used. However, in contrast to the impedance results,

TABLE 2

Complex least squares fitting of the synthetic admittance data associated with the impedance plot of Fig. 2 when error is introduced by digit roundoff

W	Sector	n	R_1/Ω	R_2/Ω	$C_2/\mu\text{F}$
0	1	4	9.96 ± 0.22	1000.04 ± 0.21	1.0000 ± 0.0004
0	1	3	16.67 ± 1.87	993.59 ± 1.83	1.0133 ± 0.0038
P	1	3	16.83 ± 8.11	993.43 ± 7.97	1.0137 ± 0.0165
0	2	3	10.39 ± 0.40	1000.09 ± 0.31	1.0014 ± 0.0008
0	3	3	10.00 ± 0.001	1002.30 ± 3.20	1.0001 ± 0.0002
0	4	3	10.00 ± 0.005	1000.50 ± 0.61	1.0005 ± 0.0001
0	2	2	2.54 ± 2.92	1002.15 ± 2.20	0.9834 ± 0.0062
0	3	2	10.01 ± 0.01	979 ± 29	0.9955 ± 0.0017
0	4	2	10.00 ± 0.005	1000.75 ± 6.15	1.0002 ± 0.0010
P	4	2	10.00 ± 0.05	1004.07 ± 1.72	1.0003 ± 0.0015

Sectors 3 and 4 give better estimates, at least for R_1 and C_2 , than does Sector 2. Furthermore, the advantages, if any, of $W=P$ weighting are small. A direct comparison of Tables 1 and 2 reveals that in Sector 1 a Y analysis gives better results than the corresponding Z fit, though with $W=P$ the Y and Z results for $n=3$ are comparable. The Sector 2 Z fit (which as already noted seems to be the optimal range for Z) gives better results, especially for $n=2$, than the corresponding Y analysis. In Sector 3 the Y data give better values for R_1 , but not for R_2 and C_2 . Finally, the Y fit of Sector 4 even for the worst case of $n=2$ gives parameter values accurate to within 0.5%.

In summary, for this type of Voigt network, one can conclude from our work that with data taken over the entire frequency range a complex least squares fit of Y data gives the best results. However, for data in the middle of the impedance semicircle, a Z fit would be somewhat preferred.

One could well argue that digit round off is not really a simulation of real experimental error. As an alternative procedure we contaminated the exact (13 digit) results with random noise. Let χ represent the $2K$ -component vector of exact impedance or admittance values. Then the j th "measurement" of χ is denoted by $\tilde{\chi}^j$ with

$$\tilde{\chi}_i^j = \chi_i [1 + dG_i^j] \quad (2)$$

Here i is a component index with $1 \leq i \leq 2K$ and d is a positive number. The vector G^j has $2K$ components and contains Gaussian or normally distributed numbers having zero mean and unit standard deviation. This procedure contaminates the real and imaginary parts of the exact data completely separately so that the "errors" added to them are completely uncorrelated for each frequency considered. For large K one would expect that approximately 68% of the $\tilde{\chi}_i^j$ would have an absolute fractional deviation from their true values less than d . The G^j vectors were generated using the International Mathematical and Statistical Library (IMSL, 7th ed.) routine GGNML*. This program uses a pseudo-random number generator which for J different choices of a "seed" produces J different vectors G^j . We generated ten ($J=10$) such sets of contaminated data from the impedance values in Sector 1 ($K=25$). The results of a complex least squares fit for $d=10^{-3}$ are summarized in Table 3. The choice of GT equal to a , b , or c stands for three alternatives in analyzing the contaminated data. First from each of the ten data sets a set of parameters was determined. The average and standard deviation of these parameters over the ten trials are presented as $GT=a$. In a slightly more realistic simulation of experimental data, a new composite data set was generated with data points $\tilde{\chi}_i^c$ given by

$$\tilde{\chi}_i^c = \left[\sum_{j=1}^J \tilde{\chi}_i^j \right] / J \quad (3)$$

* IMSL is a collection of FORTRAN mathematical and statistical subroutines developed by International Mathematical and Statistical Libraries, Inc., 7500 Bellaire Boulevard, Houston, TX 77036, U.S.A.

TABLE 3

Complex least squares fitting results for the synthetic impedance data of Fig. 2 when zero mean, normally distributed random error is added to each point

<i>W</i>	Sector	<i>GT</i>	R_1/Ω	R_2/Ω	$C_2/\mu\text{F}$
0	1	<i>a</i>	11.89 ± 7.76	998.15 ± 7.30	1.0038 ± 0.0154
<i>P</i>	1	<i>a</i>	11.51 ± 6.10	998.42 ± 5.86	1.0033 ± 0.0122
0	1	<i>b</i>	10.61 ± 2.30	999.29 ± 2.25	1.0013 ± 0.0046
<i>P</i>	1	<i>b</i>	11.38 ± 1.43	998.53 ± 1.40	1.0029 ± 0.0029
<i>S</i>	1	<i>b</i>	11.15 ± 1.41	998.76 ± 1.37	1.0024 ± 0.0028
0	1	<i>c</i>	23.46 ± 10.01	986.90 ± 9.77	1.0270 ± 0.0203
<i>P</i>	1	<i>c</i>	21.16 ± 5.81	989.15 ± 5.66	1.0221 ± 0.0120
<i>S</i>	1	<i>c</i>	18.60 ± 6.52	991.60 ± 6.34	1.0168 ± 0.0131

and intrinsic uncertainty

$$\sigma_i^c = \left\{ \sum_{j=1}^J [\bar{x}_i^j - \bar{x}_i^c]^2 / (J - 1) \right\}^{1/2} \quad (4)$$

Thus, in addition to the $W=0$ and $W=P$ options, one now has the "natural" weighting choice, $W=S$, where $w_i^c = (\sigma_i^c)^{-2}$. The analysis performed on the composite data is called *b*. Finally, the $GT=c$ analysis was made on the $GT=b$ composite data rounded off to three digits. It should be emphasized that the uncertainties listed for $GT=b$ or $GT=c$ come from a least squares estimation of the parameter covariance matrix, while those for $GT=a$ are just the ordinary standard deviations associated with estimating an average based on ten trials.

As one would expect, the $GT=b$ analysis gives the best results. Furthermore, a comparison between Tables 1 and 3 shows that the effect of rounding to three digits is a more severe deterrent to an accurate determination of parameters than the presence of random noise in the third digit. It should be noted, however, that this conclusion is being made for the worst case situation of data from Sector 1. Nevertheless, from these results one may conclude that if the data are adequately distributed in frequency a complex least squares fit has very strong resolving power even for data accurate to 0.1% or worse.

Although we have not investigated the effect of adding fully correlated random error to real and imaginary data values associated with the same frequency, one should realize that round-off or truncation induces correlation between the errors thereby produced. Thus, our round-off results, which are generally worse than those with "equivalent" uncorrelated Gaussian error added, throw some light on this question. But equivalent truncation or round-off will generally yield worse data in any event since added Gaussian error will not disturb some data values much or at all, leaving them at or close to their proper values, while truncation degrades the accuracy of all points.

A more complicated and more relevant problem in parameter estimation occurs

for a Voigt network having two distinct time constants. The determination of lump circuit elements is then closely related to the problem of the resolution of two overlapping semi-circles (in this context in the impedance plane, but other applications would include nearly overlapping Debye curves for complex dielectric permittivity). There is already a great deal of literature on this subject [9,10,17-19]. These methods are based on either direct (i.e., finding the location of centers and radii) iteration or transformed (from circles to straight lines) iteration procedures. Thus a study of the resolving power of a complex least squares estimation for this problem is warranted.

Synthetic Z and Y data were generated for the following set of parameters: $(R_1, C_1; R_2, C_2) = (10^3 \Omega, 0.1 \mu\text{F}; 10^2 \Omega, 10^2 \mu\text{F})$. The data were chosen to be approximately equally spaced in log frequency with a frequency ratio of $10^{1/8}$. A total of 48 data points were generated starting at $f = 0.1$ Hz and ending at $f = 74.99$ kHz. The above choice of parameter values yields a time constant ratio of 100 and 10:1 difference in the size of the two semi-circular arcs in the resulting impedance plane plot. A three-dimensional impedance plot is shown in Fig. 3, where the origin of the $\log(f)$ scale is taken to be -1 .

As before, to simulate error the exact data values were rounded to n digits. The results of a complex least squares fit are presented in Table 4, where the choice $T = Z$ or Y indicates the type of data (Z or Y) which was analyzed. These results again show the general superiority of $W = P$ weighting compared to the unweighted $W = 0$ fit, but the difference in resolution ability between Z and Y analyses is not clear. One can, however, see that the parameter estimates are good even for $n = 2$.

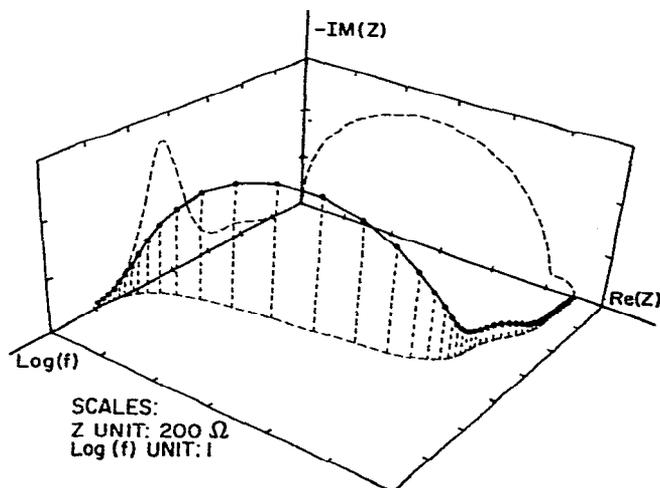


Fig. 3. Perspective three-dimensional plot of the Z^* response of a $N=2$ Voigt circuit when $(R_1, C_1; R_2, C_2) = (10^3 \Omega, 0.1 \mu\text{F}; 10^2 \Omega, 10^2 \mu\text{F})$. Two-dimensional projections are shown as dashed lines.

TABLE 4

Complex least squares fitting results for synthetic data of a Voigt circuit with two time constants having a ratio of 100

W	T	n	R_1/Ω	$C_1 \times 10/\mu\text{F}$	R_2/Ω	$C_2/\mu\text{F}$
0	Z	3	1000.04 ± 0.28	1.00015 ± 0.00071	100.28 ± 0.34	101.77 ± 0.94
P	Z	3	999.78 ± 0.19	1.00011 ± 0.00019	100.02 ± 0.06	99.992 ± 0.117
0	Y	3	1000.00 ± 1.48	1.00042 ± 0.00013	100.07 ± 2.56	100.01 ± 7.11
P	Y	3	1000.03 ± 0.16	1.00028 ± 0.00023	99.981 ± 0.061	100.11 ± 1.04
0	Z	2	996.3 ± 2.5	1.0020 ± 0.0064	106.5 ± 3.0	94.93 ± 7.35
P	Z	2	999.3 ± 1.7	1.0028 ± 0.0017	99.83 ± 0.51	99.84 ± 1.03
0	Y	2	998.0 ± 1.1	0.9955 ± 0.0009	101.7 ± 18.8	93.6 ± 48.1
P	Y	2	998.9 ± 1.2	0.9987 ± 0.0017	100.29 ± 0.45	99.25 ± 0.76

Four other sets of synthetic data with different ratios of time constants and arc size were generated. From the point of view of parameter resolution the most stringent case had $(R_1, C_1; R_2, C_2) = (10^3 \Omega, 20 \mu\text{F}; 10 \Omega, 10^3 \mu\text{F})$. These choices lead to a time constant ratio of 2 and a size ratio of 200:1. The 32 points used here began at $f = 0.1$ Hz and extended to $f = 749.9$ Hz, with approximately equal spacing in log frequency (frequency ratio again equal to $10^{1/8}$). Table 5 shows the results of a complex least squares analysis of these data. For $n < 4$ and $W = 0$ the results for R_2 and C_2 are wholly uncertain and therefore are not shown; this is also essentially the case even for $W = P$ when $n = 2$.

Again one sees that the $W = P$ weighting option results in the best estimates. Although, in contrast to the conclusions drawn from Table 4, the choice of analyzing Y rather than Z data here gives more accurate parameter values. Finally, as one would expect for this very stringent case of nearly equal time constants and one arc nearly negligible in size compared to the other, the inherent accuracy of the input Z or Y data must be greater than was the case of the data of Table 4 in order to achieve adequate resolution. However, as the $n = 3$, $W = P$, Y fitting results illustrate, one can still resolve all parameters (and in particular R_2 and C_2) to within less than a single estimated standard error of their "correct" values. In summary,

TABLE 5

Complex least squares fitting results for synthetic data of a Voigt circuit with two time constants having a ratio of 2

W	T	n	R_1/Ω	$C_1/\mu\text{F}$	R_2/Ω	$C_2/\mu\text{F}$
0	Z	4	1001.3 ± 1.6	19.961 ± 0.045	8.6 ± 1.6	1111 ± 131
P	Z	4	1000.39 ± 0.44	19.989 ± 0.013	9.58 ± 0.43	1031 ± 34
0	Y	4	999.91 ± 0.17	20.00 ± 0.12	10.0 ± 4.1	999 ± 304
P	Y	4	999.96 ± 0.23	20.00 ± 0.01	10.04 ± 0.23	998 ± 17
P	Z	3	1005.7 ± 1.6	19.83 ± 0.05	4.38 ± 1.47	1787 ± 404
P	Y	3	1001.8 ± 1.7	19.96 ± 0.05	8.35 ± 1.73	1130 ± 167

these and similar synthetic data results show that in general the resolution of a complex least squares fitting can be very high provided a sufficient frequency region is covered.

III. EXPERIMENTAL DATA

The use of synthetic data in Section II had the advantage of being a completely controlled way of seeing how the introduction of error affects the ability of a complex least squares procedure to yield estimates of parameter values. For experimental measurements this is generally not the case and so the success of the method in modeling measured data taken on real systems is the ultimate criterion of its usefulness.

The frequency response of the actual $N=3$ ladder network of lumped-constant circuit elements of Fig. 4 was measured using a Solartron type 1172 response analyzer. The experimental cell arrangements and equipment have been described elsewhere [20]. Here the real values of Y had four decimal places and the imaginary parts had either three or four. Frequencies were equally spaced in $\log(f)$ with a ratio of about 1.58 and extended from 0.4 Hz to 10^4 Hz. Three-dimensional perspective plots of both the Z and Y data are presented in Figs. 5 and 6 respectively. In Fig. 5 the origin of the $\text{Re}(Z)$ axis is at 2.5 k Ω and in Fig. 6 the origin of the $\text{Re}(Y)$ axis is at 120 μS . As usual, the origins of the imaginary axes are at zero. There is very little separation apparent in Figs. 5 and 6 between the sections having 27.5 μs and 211 μs time constants, even though their ratio is 7.7. Unity weighting proved to yield results with slightly smaller parameter standard deviations for these data than did $W=P$ weighting. The parameter estimates found from Y fitting are shown in parentheses in Fig. 4.

These results show good agreement with the nominal values. It is likely that the least squares estimates are, in fact, appreciably more accurate than the nominal

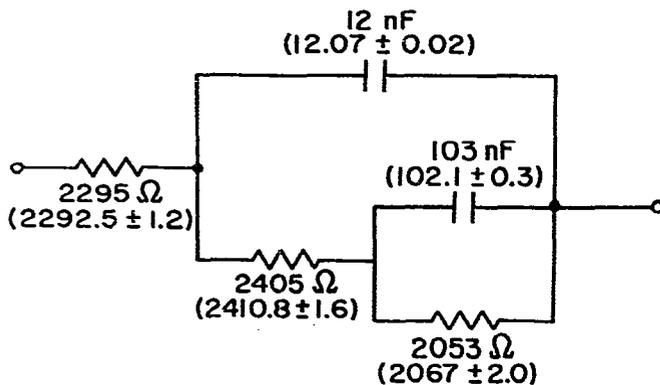


Fig. 4. Test circuit involving lumped elements. Nominal values are the numbers on top, while those in parentheses are the complex least squares estimates.

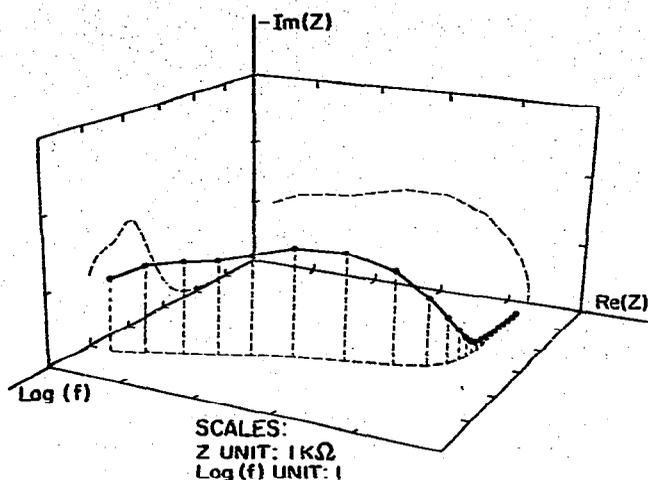


Fig. 5. Perspective three-dimensional plot of the Z^* response of the circuit shown in Fig. 4.

values since the former represent the result of many individual measurements. As a check the 12 nF capacitor was remeasured at $f = 120$ to 10^3 Hz by using a General Radio type 1680 bridge, yielding values which ranged from about 11.9 to 12.25 nF with some slight tendency towards increasing values with increasing frequency. The mean and standard deviations of nine measurements were $12.09 \pm 0.15 \text{ nF}$, in close agreement with the fitting result. Fitting to data in impedance rather than admittance form gave parameter estimates very close to those in Fig. 4 but with appreciably larger parameter standard deviations. On the scale of Figs. 5 and 6 no difference

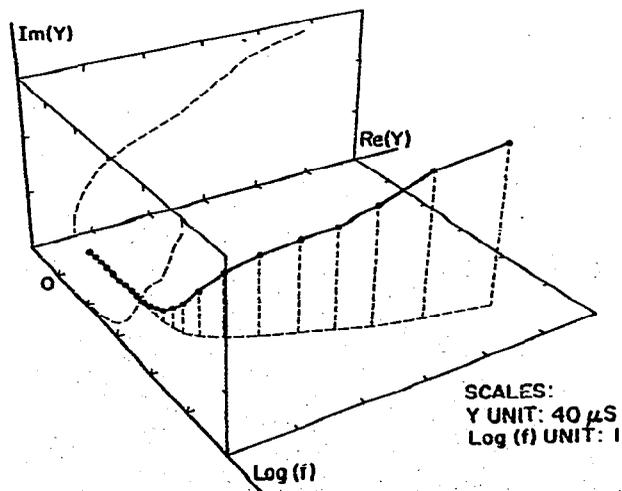


Fig. 6. Perspective three-dimensional plot of the Y response of the circuit shown in Fig. 4.

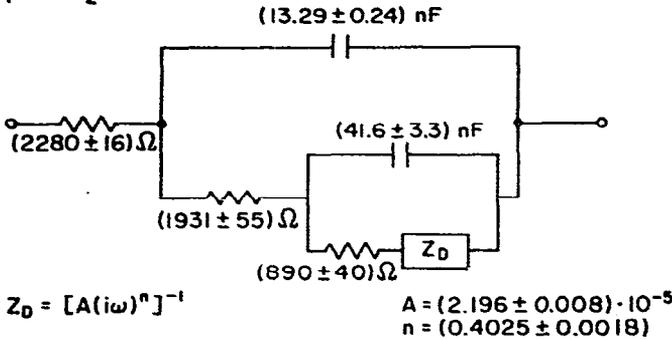
β -PbF₂ at 474 K

Fig. 7. Equivalent circuit used to fit admittance frequency data of β -PbF₂ at 474 K. Complex least squares estimates and the estimated standard deviations are shown in parentheses. A constant phase element is used here for Z_D .

would be seen between the original data and values predicted from the fit.

Finally, to illustrate the application of complex least squares to actual measurements on a distributed system, data was obtained on β -PbF₂ with platinum paint electrodes at 474 K. Both $\text{Re}(Y)$ and $\text{Im}(Y)$ were given to four decimal places and 27 points spanned the range from 0.2 Hz to 20 kHz. As usual, frequency values were taken to be exact. The circuit used for fitting is shown in Fig. 7 with Z_D given by a constant-phase element [21] of the form $Z_D = [A(i\omega)^n]^{-1}$. Here A and n are frequency independent parameters and $0 < n < 1$. A Z fit with $W = P$ gave the

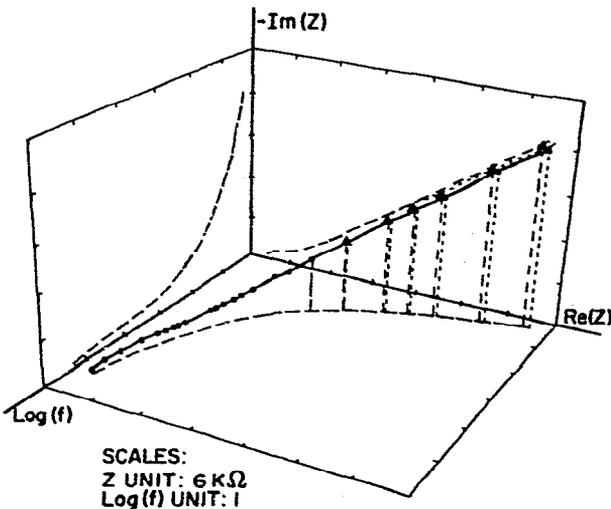


Fig. 8. Perspective three-dimensional plot of the Z^* response of β -PbF₂ at 474 K. The data points are indicated by solid circles and the complex least squares fitting results by solid triangles.

following parameter estimates: $R_\infty = (2280 \pm 16) \Omega$, $R_R = (1931 \pm 55) \Omega$, $C_R = (13.29 \pm 0.24) \text{ nF}$, $R_A = (890 \pm 40) \Omega$, $C_A = (41.6 \pm 3.3) \text{ nF}$, $A = (2.196 \pm 0.008) \times 10^{-5}$, and $n = (0.4025 \pm 0.0018)$. Despite the fact that seven parameters are used, the relative parameter uncertainty estimates are all less than about 8%.

The actual complex least squares fitting results are shown as a three-dimensional impedance plot in Fig. 8. Here the original data points are indicated by small solid circles (a solid line in three dimensions), while points at the same frequencies calculated using the fitted parameters values in the circuit are designated by small solid triangles. Most of the original and calculated points fall so close together that they cannot be distinguished. Only at the lowest frequencies do any differences (whose relative errors are still less than 2%) appear. The lines with short dashes are associated with the original data and those with longer dashes with the calculated points. These results are only one specific example of the power and utility of employing together non-linear complex least squares analysis and three-dimensional perspective plotting.

APPENDIX

Fitting data to a circle

The problem of finding the "best circle" through a set of data points in the plane (x_i, y_i) , with $i = 1, \dots, N$, is not quite as arcane or pedantic an exercise as it might at first seem. The problem has arisen in several quite applied contexts [8,22] and could be related to the problem of resolving multiple time constants [9] discussed in Section II of this paper. Also, in his text, Brandt [23] discusses the least squares solution of this problem. As necessary ingredients to such a least squares analysis one needs initial estimates of the radius, R^0 , and the coordinates of the circle's center, (P_x^0, P_y^0) . Following Brandt one could arbitrarily pick any three points (to fit points to a circle, of course one must have $N \geq 3$) and obtain the following estimates:

$$\begin{bmatrix} P_x^0 \\ P_y^0 \end{bmatrix} = 0.5 \times \begin{bmatrix} (y_3 - y_2)[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)] - (y_2 - y_1)[(x_3^2 + y_3^2) - (x_2^2 + y_2^2)] \\ (x_2 - x_1)[(x_3^2 + y_3^2) - (x_2^2 + y_2^2)] - (x_3 - x_2)[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)] \end{bmatrix} \times [(x_2 - x_1)(y_3 - y_2) - (x_3 - x_2)(y_2 - y_1)]^{-1} \quad (\text{A1})$$

$$R^0 = [(x_1 - P_x^0)^2 + (y_1 - P_y^0)^2]^{1/2} \quad (\text{A2})$$

Here the points are labeled by 1, 2 and 3 for simplicity. The problem with this procedure is that it does not take all N points into account, and hence the reliability of eqns. (A1) and (A2) is very dependent upon the initial choice of the three points.

One could of course consider all $N!/3!(N-3)!$ triplets and average the results of eqns. (A1) and (A2). However, this could be rather time consuming and is also unduly influenced by the effects of a few points having large deviations from the "best circle". As an alternative approach we give here a generalized—inverse or least squares—solution [24] of the over-determined system of equations,

$$W^{1/2}AP = W^{1/2}B \quad (\text{A3})$$

Here A is an $N \times 3$ matrix, whose i th row has the form

$$A_i = (x_i, y_i, 1) \quad (\text{A4})$$

$W^{1/2}$ is an $N \times N$ weighting matrix,

$$W_{ij}^{1/2} = (w_i)^{1/2} \delta_{ij} \quad (\text{A5})$$

where w_i is the weight associated with the point (x_i, y_i) . Finally B is a $N \times 1$ column vector with

$$B_i = -(x_i^2 + y_i^2) \quad (\text{A6})$$

while P is a 3×1 column vector with

$$P = \begin{bmatrix} -2P_x \\ -2P_y \\ C \end{bmatrix} \quad (\text{A7})$$

The parameter C is related to the other parameters by

$$C = (P_x)^2 + (P_y)^2 - (R)^2 \quad (\text{A8})$$

For more than three points not all on the same circle, eqn. (A3) has no solution. However, if at least three of the points are not co-linear, then the matrix A has full column rank. Then the generalized solution of eqn. (A3), which is defined as the unique minimum of the function $f_1(P_x, P_y, R) \equiv (AP - B)^T W (AP - B)$, is given by

$$P^0 = (A^T W A)^{-1} A^T W B \quad (\text{A9})$$

where $W = W^{1/2} W^{1/2}$ and A^T represents the transpose of A .

This method has the weakness that each point (x_i, y_i) is assumed to have an a priori weight, w_i , which is a measure of how well that point lies on a circle. These weights are not easily related to any intrinsic uncertainties in the measurement of the coordinates x_i and y_i . In particular, no provision is made here for the case where the error in the x measurements is appreciably different from the error in measuring y . However, such intrinsic x and y weights would be used in the actual least squares calculation of P_x , P_y , and R for which P_x^0 , P_y^0 and R^0 are intended only as good starting estimates. In the absence of any other information, the choice $w_i = 1$ would probably suffice. If one defines the average of a quantity a as

$$\langle a \rangle \equiv \left[\sum_{j=1}^N a_j w_j \right] / \left[\sum_{j=1}^N w_j \right] \quad (\text{A10})$$

then eqn. (A9) becomes

$$P^0 = - \begin{bmatrix} \langle x^2 \rangle & \langle xy \rangle & \langle x \rangle \\ \langle xy \rangle & \langle y^2 \rangle & \langle y \rangle \\ \langle x \rangle & \langle y \rangle & 1 \end{bmatrix}^{-1} \begin{bmatrix} \langle x(x^2 + y^2) \rangle \\ \langle y(x^2 + y^2) \rangle \\ \langle x^2 + y^2 \rangle \end{bmatrix} \quad (\text{A11})$$

Hence, one gets the following results:

$$P_x^0 = [\text{cov}(y, x^2 + y^2) \times \text{cov}(x, y) - \text{cov}(x, x^2 + y^2) \times \sigma^2(y)] / D \quad (\text{A12})$$

$$P_y^0 = [\text{cov}(x, x^2 + y^2) \times \text{cov}(x, y) - \text{cov}(y, x^2 + y^2) \times \sigma^2(x)] / D \quad (\text{A13})$$

$$R^0 = \left[(P_x^0)^2 + (P_y^0)^2 + \langle x^2 + y^2 \rangle - 2P_x^0 \langle x \rangle - 2P_y^0 \langle y \rangle \right]^{1/2} \quad (\text{A14})$$

Here the covariance of quantities a and b [25] is denoted by

$$\text{cov}(a, b) \equiv \langle (a - \langle a \rangle)(b - \langle b \rangle) \rangle = \langle ab \rangle - \langle a \rangle \langle b \rangle \quad (\text{A15})$$

while the variance of a is defined as

$$\sigma^2(a) \equiv \text{cov}(a, a) \quad (\text{A16})$$

The denominator in eqns. (A12) and (A13) is given by

$$D = 2 \left\{ [\text{cov}(x, y)]^2 - \sigma^2(x) \times \sigma^2(y) \right\} \quad (\text{A17})$$

At this point it should again be emphasized that eqns. (A12) through (A14) are not the least squares solution of the problem of fitting points to a circle. That is an inherently non-linear problem in contrast to the linear result of eqn. (A11). In fact, it is not difficult to show that P_x^0 , P_y^0 and R^0 minimize the sum

$$f_1(P_x, P_y, R) = \sum_{i=1}^N w_i [(x_i - P_x)^2 + (y_i - P_y)^2 - R^2]^2 \quad (\text{A18})$$

In contrast, for unit weighting assigned to both x and y values, a least squares analysis gives values of P_x , P_y , and R which minimize the sum

$$f_2(P_x, P_y, R) \equiv \sum_{i=1}^N \left\{ [(x_i - P_x)^2 + (y_i - P_y)^2]^{1/2} - R \right\}^2 \quad (\text{A19})$$

If one is interested in the less general problem where one of the coordinates of the circle's center, say P_y , is known to be zero, then the parameters P_x^0 and R^0 which minimize the function $f_1(P_x, P_y \equiv 0, R)$ are given by

$$P_x^0 = 0.5 \times [\text{cov}(x, x^2 + y^2)] / [\sigma^2(x)] \quad (\text{A20})$$

$$R^0 = [\langle x^2 + y^2 \rangle + P_x^{02} - 2\langle x \rangle P_x^0]^{1/2} \quad (\text{A21})$$

if one exchanges x and y , then eqns. (A20) and (A21) give the results for the case where P_x is fixed at 0.

To connect these results to the determination of model circuit parameters, consider the Voigt network of Section II again with $(R_1, C_1; R_2, C_2) = (10 \Omega, 0; 1000 \Omega, 1 \mu\text{F})$. From a knowledge of the problem one has $P_y^0 = 0$. Hence, one deduces that $R_1 = P_x^0 - R^0$ and $R_2 = 2R^0$. Following Despic [8], the value of the capacitor C_2 can be found independently of R_1 and R_2 as the slope of $-\omega/\text{Im}(Z)$ vs. ω^2 . This result can be found by the non-weighted linear least squares formula, $C_2 = \text{cov}[\omega^2, -\omega/\text{Im}(z)]/\sigma^2(\omega^2)$. Of course this result has used only half of the data available by ignoring the real component of the frequency response. In addition, one has no parameter uncertainty estimates from this approach. However, it is interesting to compare the results obtained by this simple analysis with the nonlinear complex least squares determination of the parameters. Using the synthetic data of Sector 4 with $n=2$, one obtains the estimates $R_1 = 9.66 \Omega$, $R_2 = 998.34 \Omega$, and $C_2 = 1.0029 \mu\text{F}$. One sees that these results, particularly those for R_1 and R_2 which were determined from the circle fitting, are extremely close to the corresponding $W=0$ results of table 1. Thus, it would appear that eqns. (A20) and (A21) provide good initial estimates to an actual least squares analysis.

ACKNOWLEDGEMENTS

We are grateful to Mr. John Dale for writing and implementing the three-dimensional perspective plotting routines. One of us (A.P.L.) would like to thank Dr. W.R. Mann of the Department of Mathematics of the University of North Carolina, Chapel Hill for many helpful discussions on least squares problems. This work was supported in part by U.S. National Science Foundation Grant DMR 80-05236 and by NATO under Research Grant 1696.

REFERENCES

- 1 J.H. Sluyters, *Recl. Trav. Chim. Pays-Bas*, 79 (1960) 1092.
- 2 A recent review by W.I. Archer and R.D. Armstrong appears in *Specialist Periodical Reports—Electrochemistry*, Vol. 7, The Chemical Society, London, 1980, pp. 157–202.
- 3 D.R. Powell and J.R. Macdonald, *Comput. J.*, 15 (1972) 148.
- 4 J.R. Macdonald and J.A. Garber, *J. Electrochem. Soc.*, 124 (1977) 1022.
- 5 S. Brandt in *Statistical and Computational Methods in Data Analysis*, North Holland, Amsterdam, 1970, Chapter 9.
- 6 H.I. Britt and R.H. Luecke, *Technometrics*, 15 (1973) 233.
- 7 D.D. Macdonald, *J. Electrochem. Soc.*, 125 (1978) 2062 and discussions by D.D. Macdonald, J.R. Macdonald and D.R. Franceschetti in *J. Electrochem. Soc.*, 126 (1979) 1082.
- 8 A.R. Despic, V.D. Jovic and M.V. Djorovic, *Israel J. of Chem.*, 18 (1979) 65.
- 9 M. Kleitz and J.H. Kennedy in P. Vashishta, J. Mundy and G. Shenoy (Eds.), *Fast Ion Transport in Solids*, Elsevier, North Holland, New York, 1979, pp. 185–188.
- 10 T.M. Gur, I.D. Raistrick and R.A. Huggins, *Solid State Ionics*, 1 (1980) 251.
- 11 J.R. Macdonald in 1980 Annual Report, 49th Conference on Electrical Insulation and Dielectric Phenomena, National Academy of Sciences, Washington, D.C., 1980, pp. 3–49. Reprinted with corrections in *IEEE Trans. Electr. Insul.*, EI-15 (April, 1981) 65.
- 12 J.R. Macdonald in M. Kleitz and J. Dupuy (Eds.), *Electrode Processes in Solid State Ionics*, D. Reidel, Dordrecht, Holland, 1976, pp. 149–183.

- 13 I.M. Novosel'skii, N.N. Gudina and Yu.I. Fetistov, *Sov. Electrochem.*, 8 (1972) 546.
- 14 Ref. 5, pp. 170-173.
- 15 E.W. Marquardt, *SIAM J. Appl. Math.*, 11 (1963) 431.
- 16 J.J. Moré in G.A. Watson (Ed.), *Lecture Notes in Mathematics*, Vol. 630, Numerical Analysis, Springer Verlag, Berlin, 1978, pp. 105-116.
- 17 A. von Hippel, D.B. Knoll, W.B. Westphal, M.A. Maidique, R. Mykolajewycz and J. Iglesias, M.I.T. Lab. Insul. Res., Tech. Rep. 6, Boston (1969) 53.
- 18 A. von Hippel, D.B. Knoll and W.B. Westphal, *J. Chem. Phys.*, 54 (1971) 134.
- 19 G.W. Gross, I.C. Havslip and R.N. Hoy, *Geophysics*, 45 (1980) 914.
- 20 J. Schoonman, L.J. Stijl, J.R. Macdonald and D.R. Franceschetti, *Solid State Ionics*, 3/4 (1981) 365.
- 21 P.H. Bottelberghs and G.H.J. Broers, *J. Electroanal. Chem.*, 67 (1975) 155.
- 22 N.E. Hill, *J. Phys. C*, 13 (1980) 6273.
- 23 Ref. 5, pp. 194-197.
- 24 B. Noble, *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, 1969, pp. 39-43 and pp. 142-145.
- 25 Ref. 5, p. 22.