

IMPEDANCE SPECTROSCOPY: OLD PROBLEMS AND NEW DEVELOPMENTS

JAMES ROSS MACDONALD

Department of Physics and Astronomy, University of North Carolina, Chapel Hill,
NC 27599-3255, U.S.A.

(Received for publication 17 April 1990)

Abstract—The generality, scope, and limitations of Impedance Spectroscopy (IS) are discussed, with emphasis on unsupported conditions in ionic systems. For such conditions, the maximum reaction rate which can be determined from IS data is limited. The finite-length-Warburg diffusion frequency responses of unsupported and supported situations are simplified and compared, and similarities and differences emphasized. Two types of ambiguity possibly present in fitting IS data to equivalent circuits are discussed, one intrinsic and the other associated with distributed circuit elements. Powerful new features have been added to the author's complex nonlinear least squares (CNLS) fitting program, and the results of a Monte Carlo simulation study of bias and statistical uncertainty in CNLS fitting of equivalent circuit data are discussed. The program now incorporates new variable weighting choices which can greatly minimize such bias. It also allows two unknown weighting parameters of the error variance model to be automatically estimated during the least squares fitting, thus best matching the weighting to the data and yielding most appropriate estimates of the parameters of the fitting model.

Key words: impedance spectroscopy, diffusion, reaction rate, complex nonlinear least squares fitting.

INTRODUCTION

Although a number of reviews exist of the burgeoning field of Impedance Spectroscopy (IS), many of them are primarily concerned with supported situations, those where a high concentration of indifferent electrolyte is present, rather than with unsupported ones[1–13]. Although applicability only to supported situations is not always explicitly stated in these works, this restriction may often be identified by their concentration on liquid electrolytes and their assumption of electroneutrality, rather than their use of Poisson's equation. But solid materials, and even some liquid electrolyte situations of interest, are not supported and their analysis requires satisfaction of the Poisson equation throughout the material.

It is convenient to partition IS into two sub-categories, Electrochemical IS (EIS) and everything else. EIS deals with materials for which ionic conduction predominates, includes both supported and unsupported situations, and may involve either ionic motion and/or ion-vacancy motion. Besides liquid electrolytes, other ion-containing systems, such as superionic materials, non-stoichiometric ionically bonded single crystals, ionically conducting glasses and polymers, and fused salts, may also be included in this category. But it is worth emphasizing that IS, including its measurement and analysis methods, applies to other types of materials as well. In particular, it applies to materials exhibiting predominantly electronic conduction, such as single-crystal and amorphous semiconductors and polymers, and to solid and liquid dielectrics, whose electrical characteristics are associated with dipolar rotation. Now obviously most of these materials are unsupported, and in the whole area to which IS is applicable, there are

many more unsupported than supported situations. But, as mentioned above, the distinction between these possibilities is not always made clear, and one often finds equations and equivalent circuits which were derived and used for supported conditions also used without comment for the analysis of unsupported (usually solid) materials. One aim of the present work is to provide a brief history of theoretical analyses of unsupported situations and to compare some supported and unsupported equations and predictions.

In addition, the problem of two types of ambiguity in IS analysis will be discussed, and much attention given to complex nonlinear least squares (CNLS) analysis of small-signal *ac* data. Important improvements in the author's CNLS fitting program will be described, including a new and powerful method of automatic weighting choice, and representative bias in CNLS parameter estimates, derived from an extensive Monte Carlo CNLS simulation study, will be illustrated and discussed.

First, however, it is worthwhile to dispose of several minor points of usage. In recent years, it has become relatively common for writers in the IS field to refer to a "Nyquist diagram", taken to mean a plot of the values of the real and imaginary parts of a complex quantity, such as impedance, in the complex plane. Such usage should be discouraged. A quantity such as impedance is basically derived from measurements of current and potential at a single (input) port. But Nyquist's work[14] dealt with two-port measurements of feedback in amplifiers and involved input and output voltage determinations. Thus, a complex-plane Nyquist plot is intrinsically quite different from an impedance or admittance plot in this plane. Instead of "Nyquist diagram of impedance response,"

a better description would be, "complex-plane impedance plot."

Further, there is little reason to refer to "ac impedance" rather than just "impedance." Although one might define a higher harmonic impedance or even an indicial transient response impedance, these are uncommon in the IS area, and "impedance" may be taken, in its standard definition, to mean the quotient of vector voltage and vector current calculated from small-signal sinusoidal *ac* measurements. Any other type of impedance would require adjectival qualification. Also, impedance defined as above (see also Ref.[12], pp. 5, 134 and 169) needs no qualifier as "complex." Impedance is complex by definition, and when it is only real it is called resistance. As usual, let "imittance" mean any of the four quantities of major use in IS: impedance, Z ; admittance, Y ; complex modulus, M ; and complex dielectric constant, ϵ or ϵ^* . It will be denoted by $I = I' + iI''$. Proper usage should distinguish between the meaning of a superscript asterisk as indicating a complex conjugate quantity (as hereafter) or just a complex quantity. Because of this ambiguity, its use for the latter purpose should be avoided when possible. It is desirable, however, when no asterisk is employed, to use the "complex" modifier for M and ϵ to avoid ambiguity. Clearly, IS can, most properly, stand for "Immittance Spectroscopy", rather than the less general "Impedance Spectroscopy." Although it is customary in the IS field to plot $-Z''$ on the imaginary axis *vs* Z' on the real axis and term the result an impedance plane (or impedance complex plane) plot, rigorous usage, which seems excessive, requires the designation "complex conjugate impedance plot" or Z^* plot. Finally, note that presentation of IS data in Y , M , and ϵ form, as well as Z , either as complex plane plots or, even better, as 3-D perspective plots, can often yield improved resolution (see Fig. 1) and/or highlight errors in the data not otherwise apparent ([12], pp. 174-179; [13], pp. 28-31).

ANALYSIS OF UNSUPPORTED AND SUPPORTED SITUATIONS

General background

Theoretical analysis of the small-signal *ac* response of unsupported materials essentially began with the work of Jaffé[15] and Chang and Jaffé[16]. Because of the complexity of the equations governing such response for a material with charges of both signs possibly mobile, able to recombine, and possibly partly or completely blocked at an electrode, the following short precis of work in the area shows that it progressed over a period of some 25 years through approaches which, until the last, involved various approximations, simplifications, and special cases for ease of calculation.

Chang and Jaffé's failure to ensure full satisfaction of Poisson's equation was corrected in Ref.[17], which dealt with completely blocking electrodes, with positive and negative charges of equal valence numbers, z_1 and z_2 , with arbitrary diffusion coefficients, D_1 and D_2 , and included dissociation-recombination possibilities. Soon thereafter, Friauf[18] presented a

similar treatment which involved partially blocking electrodes using Chang-Jaffé (CJ) boundary conditions, often appropriate for solid materials and even useful for many liquid material situations. This work and most of that discussed below thus best applies in the EIS area to materials with parent-ion or completely blocking electrodes, not directly to those with redox reactions. Several physical situations for which the assumption of parent-ion electrodes is pertinent are discussed by Buck[19]. Beaumont and Jacobs[20] later investigated the response of a partly blocking system with charge of only a single sign mobile, results thus only strictly applicable to solids. Next came the first accurate treatment of the *ac* response of a fully dissociated, completely blocking situation with arbitrary z_i 's and D_i 's[21]. Reference [22] was the first to show finite length Warburg (FLW) diffusion response for an unsupported situation with equal D_i 's and equal z_i 's and with charge of a single sign free to discharge at an electrode. General FLW response, at the impedance level, is of the form:

$$Z_w(\omega) = Z_w(0)[\tanh[iA]^{0.5}/[iA]^{0.5}], \quad (1)$$

where A is given by $(l/l_d)^2$. Here, l is the separation between identical plane, parallel electrodes, and l_d is the diffusion length, proportional to $\omega^{-0.5}$ (see below).

Further work on FLW and other response possibilities appeared in Refs[23, 24]. Reference[3] pointed out that the main physical processes possibly important in general IS response are, for either supported or unsupported conditions, bulk, electrode reaction (including discharge of ions at parent-ion electrodes), specific adsorption, generation-recombination, and diffusion effects. When these processes are loosely coupled and thus appear in different frequency regions, it was noted that they each lead to a separate arc when overall impedance is plotted in the complex plane. The first four arcs, but not that associated with diffusion, were shown as semicircles with their centers on the real axis, although their centers may be displaced below this axis if there is a distribution of relaxation times present for the process involved ([12], pp. 16-17, 34-36, 87-94; [24]). These conclusions turned out to be false in one respect. Generation-recombination effects do not, in fact, lead to semicircular response in the Z plane. Recent work for general unsupported conditions[25] shows, however, that they do lead to all sorts of multiple arc shapes when plotted in the complex dielectric constant (or complex capacitance) plane for a completely blocking situation.

Ambiguous circuit and element response

If one were able to analyze IS data with an appropriate mathematical model derived directly from discrete microscopic analysis, there would be little or no ambiguity present. But this is rarely if ever possible for unsupported-situation data, and fitting to an equivalent circuit, preferably by CNLS, is the remaining option. For such fitting there are two possible types of ambiguity. The first arises because there are situations where different geometric connections of the (ideal) circuit elements of a fitting equivalent circuit can yield, with appropriate element values, the same impedance over all frequencies. The

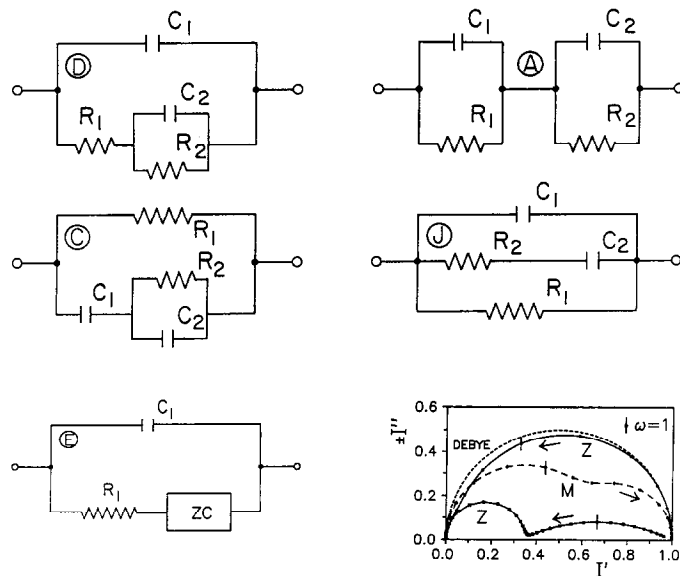


Fig. 1. Four possible ambiguous circuits (D, A, C, J), a circuit containing a distributed element (E), and comparison of some of their complex plane response curves with Debye response for the parameter values listed in Table 1. The top Z and M curves apply for D, A, C, and J, and the bottom one for the E circuit normalized to yield $I_{max} = 1$.

matter is discussed in the literature ([12], pp. 95–99; [13, 26, 27]). All equivalent two-time-constant RC circuits of this type which allow some dc conduction are shown in Fig. 1 (designated D, A, C, and J), and they and the E circuit will be discussed later. Ambiguous circuits involving inductance are also discussed in[26].

A considerable CNLS study of the theoretical response of an unsupported system was carried out in[26], and it was found that the circuit of Fig. 2, with Z_w and $R_{R\infty}$ omitted and Z_{De} taken as a FLW, was most appropriate. For this circuit, C_∞ and R_∞ account for bulk effects, C_R and R_R for electrode reaction effects, and C_A and R_A for adsorption–reaction processes. Now it is found both experimentally and theoretically that the R_A – C_A adsorption–reaction arc may involve either capacitive or (apparent) inductive effects: that is, it may appear above the real axis in a Z^* -plane plot or below it. Although an inductive element has been used to model such below-axis response, it should not

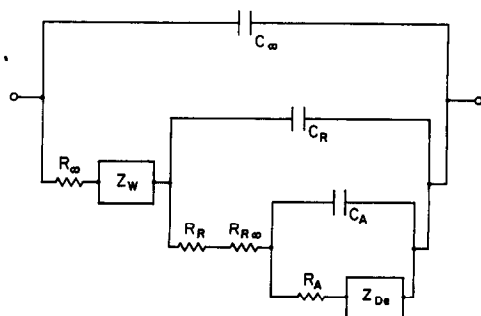


Fig. 2. General, approximate equivalent circuit for unsupported conditions. The element Z_w , when present, may account for diffusion of uncharged species in the electrodes, and Z_{De} will be zero when charge of only one sign is mobile.

be interpreted as the usual real inductance, one which involves circulating currents and energy storage in magnetic fields, but instead as a pseudo-inductance: something which yields the needed phase shift.

Further, an inductive-type phase shift associated with a below-axis arc can also be produced by a negative differential resistance in parallel with a negative differential capacitance[26]. Thus, instead of using a pseudo-inductance of large value, such negative RC elements can yield exactly the same frequency response. Although it is entirely a matter of taste which approach to use, and either is as physically reasonable as the other, I find it preferable to make the latter choice in the interest of maintaining continuity. The R_A and C_A are often positive (arc above the axis), but as adsorption rates change the arc may move below the axis. By allowing negative as well as positive values for these differential RC elements, such response can be represented by R_A and C_A under all conditions, rather than requiring a change from $R_A > 0$ and $C_A > 0$ to the use of a positive resistance and a pseudo-inductance for below-axis response. In the theory of specific adsorption presented in[26], the values of R_A and C_A indeed change from positive to negative as adsorption conditions change. This paper also showed that expressions for R_R and C_A are independent of whether CJ or the more realistic Butler–Volmer (BV) boundary conditions, which take overpotential into account, were used. In addition, a useful transformation method was described which allows response theories using the simpler CJ conditions to be converted to ones appropriate for BV conditions. Applications of the method for treating unsupported diffuse and compact double layer effects appear in[27] and[28].

In a single set of IS frequency response measurements, there is no way to avoid intrinsic circuit ambiguity when it is present. Then why does it matter? It matters because one particular geometric

arrangement of the circuit elements is more likely to model better the actual connectivity of the physical processes present than are the other possibilities. Expressions for the circuit elements in terms of microscopic quantities will be simpler for this most appropriate arrangement and will generally show simplest and most physically reasonable dependencies on temperature, potential, and electrode spacing, l . It is thus clear that if measurements are carried out for a range of temperatures, potentials, and/or l values and are fitted using CNLS to the various equivalent circuits, then this type of ambiguity may be resolved.

It is frequently impossible, unfortunately, to obtain a reasonable fit of IS data to an equivalent circuit which involves only a small number of R 's and C 's. The vast majority of such data involve at least one appreciable frequency region where impedance or admittance shows fractional frequency response and is thus proportional to $\omega^{\pm n}$, where $0 \leq n \leq 1$. In the absence of an exact response solution at the microscopic level, behavior of this kind is best handled by using an equivalent circuit containing one or more distributed circuit elements (DCE's). Such elements include, for example, FLW response, more general diffusion response, the constant phase element (CPE), Havriliak–Negami response, Williams–Watts (WW) response, and distribution of activation energies (DAE) response[12, 13, 29, 30]. The necessity of using equivalent circuits containing such elements leads to the second kind of ambiguity present in the fitting of IS data. It turns out, at least to first order, that any DCE which leads to symmetrical response in the impedance plane, such as Cole–Cole (denoted ZC for conductive systems), can be well fitted by any other symmetrical DCE. Further, nearly any DCE which leads to unsymmetrical response can be at least reasonably well fitted by any other such DCE[12, 13, 29, 31–35]. Space limitations forbid illustrations of these ambiguities herein. They are not intrinsic, however. For a response region which is well separated from those of other processes and with sufficiently accurate data and/or data which extend well away from peak response in both frequency directions, CNLS fitting does allow adequate discrimination to be made, provided appropriate weighting is used. See especially the results described in Ref.[35]. Of the various DCEs available, the ones which can best fit the others and a wide range of experimental data involving thermal activation are the DAEs[29, 31–35]. They further have the virtue of providing specific and physically meaningful predictions for the temperature dependence of the fractional exponent n or its equivalent.

Recently, much effort has been devoted to developing fractal theories of rough electrode response, eg [36–39], to explain CPE-like behavior, $Z \propto (i\omega)^{-n}$, for either supported or unsupported conditions. Although there may indeed be instances where electrode-interface response is primarily associated with fractal structures, they may be rare. Consider the following. First, it seems somewhat unlikely that typical electrode roughness and pores could involve the necessary self-similarity over more than three to four levels, thus not leading to a good approximation to full self-similar behavior. Second, there are numerous other theories which lead to CPE response[29],

including many physically based ones which are associated with such response in solids. Thus, before ascribing CPE behavior to fractal structures, one should first establish that this behavior is indeed associated with intensive electrode-interface regions. But even when this has been verified, the matter remains unproven. Bates *et al.*[40] have carefully compared measured CPE response for electrode-interface regions with well characterized electrode roughness profiles and found no correlation between the CPE fractional exponent n and the fractal dimension of the rough interface. An approach possibly preferable to the fractal one which may still yield CPE response may involve a detailed analysis of the effect of pore shapes and size-shape distributions, eg [38]. It seems likely that analysis of deviations from exact CPE response will help allow discrimination between the many various theories which lead, at least approximately, to such response over a limited frequency range.

Some differences in unsupported and supported response

The exact continuum treatment of small-signal IS response of unsupported situations without dc bias culminated in the work of Ref.[41]. Mobile positive and negative charges were assumed present having arbitrary D_i 's and z_i 's and general CJ boundary conditions were used. The charges were taken to arise from intrinsic dissociation and/or from extrinsic donor or acceptor centers. Arbitrary dissociation–recombination parameters appear for all three types of charge generation. The results are appropriate for either solid or liquid materials between plane, parallel electrodes and apply for either ionic or electronic conduction. All five physical processes discussed above in reference to [3] contributed to the overall response. Because of the complexity of the situation and its solution, not all the implications of the latter are even yet fully clear. But the exact response for completely blocking conditions has recently been examined in detail[25]. A few further results are discussed below.

The theoretical work of [41] showed that the circuit of Fig. 2 with Z_w , $R_{R\infty}$, and Z_{De} omitted was virtually exactly applicable for the common fully dissociated intrinsic situation with charge of only one sign mobile and free to discharge at the electrodes. This hierarchical-response circuit has been widely used for both supported and unsupported situation data fitting. Under the above conditions, no FLW response associated with the mobile charge appears, but such response may still arise from diffusion of uncharged (discharged) entities in the material or the electrodes[42, 43]. But later CNLS fitting of the exact response for arbitrary D_i values, but with charge of one sign completely blocked showed[44] that in the absence of specific adsorption (and presumably in its presence as well) if R_R is the usual reaction resistance, then an additional non-zero, non-reaction resistance, $R_{R\infty}$, is necessary, as well as a possible FLW response element for Z_{De} , for unsupported situations. Thus $R_{R\infty}$ is a necessary element of the circuit even in the limit of infinite discharge rate, where $R_R = 0$. Its presence in the present ambipolar diffusion situation arises from the drag of charges of one sign on

those of the other sign. Since $R_{R\infty}$ was found to be proportional to L_D/l , the larger the mobile charge concentration and the larger the electrode separation the smaller it will be. Here, L_D is the Debye length. Nevertheless, for fixed conditions it sets an upper limit for accurate estimation of the reaction rate constant from IS measurements on unsupported systems, a limit which decreases proportionally as the mobility of the reacting charge decreases[44].

Another important difference between supported and unsupported IS response appears when diffusion effects are important. The exact solution in the unsupported case[24, 26, 41] shows that a diffusion arc may appear when charges of both sign are mobile and at least one or the other of them is partly or completely blocked at the electrodes. It is not present for completely ohmic electrodes or for complete blocking of charges of both signs[25]. The expression for the A of equation (1) is, in the notation of these papers, $A = M^2 b \omega R_\infty C_\infty$, which becomes on expansion[41]:

$$A = (\omega l^2 e / 4kT) \left[\frac{z_1^2 p_e + z_2^2 n_e}{z_1 \mu_1 p_e + z_2 \mu_2 n_e} \right] \times \left[\frac{(\mu_1 / \mu_2) + 2 + (\mu_2 / \mu_1)}{(z_1 / z_2) + 2 + (z_2 / z_1)} \right], \quad (2)$$

where p_e and n_e are equilibrium values of the mobile charge and μ_1 and μ_2 are their mobilities, respectively. The quotient e/k may be replaced by F/R , the Faraday and gas constants. Now the use of the bulk electroneutrality condition, $z_1 p_e = z_2 n_e$, and the Nernst-Einstein relation, $D_i = (kT/e)(\mu_i/z_i)$, leads to:

$$A = (\omega l^2 / 4) \left[\frac{z_1 + z_2}{D_1 z_1 + D_2 z_2} \right] \times \left[\frac{(D_1 z_1 / D_2 z_2) + 2 + (D_2 z_2 / D_1 z_1)}{(z_1 / z_2) + 2 + (z_2 / z_1)} \right]. \quad (3)$$

On factoring this equation, one obtains the new simplified symmetrical results:

$$A = (\omega l^2 / 4) \left[\frac{D_1 z_1 + D_2 z_2}{(D_1 D_2)(z_1 + z_2)} \right] = (\omega l^2 e / 4kT) \left[\frac{\mu_1^{-1} + \mu_2^{-1}}{z_1^{-1} + z_2^{-1}} \right] \equiv [l/l_d]^2, \quad (4)$$

appropriate for use in equation (1) for sufficiently low concentrations that one need not introduce activity coefficients.

For the supported case, FLW results which are consistent with equation (1) seem first to have been derived by Labes and Lullies[45-47] in the biological membrane area. Because of the decoupling associated with the indifferent electrolyte which supplies the support, one obtains the same result for a charged (parent-ion electrode) or uncharged diffusing entity ([12], p. 105; [19, 41]). If its diffusion coefficient is D and the charged species is univalent, then $A \equiv \omega l^2 / 4D$, in agreement with equation (4) when $D_1 = D_2 \equiv D$ and $z_1 = z_2 = 1$. When there are two charged species present in a supported situation, both free to discharge, each species leads to the presence of a separate FLW contribution of the form of equation (1), as in the redox situation, where there are both reduced and oxidized mobile species in the

electrolyte ([11; 12], p. 106). Then, unsupported and supported response may be quite different.

Thus far we have dealt with response for flatband conditions: ones where, in the absence of any applied pd , the concentrations of mobile charged species are uniform throughout the region between electrodes. But even in the absence of an applied pd , a Frenkel space charge layer can be produced near an electrode because of the difference of work function between the electrode and conducting material. Further, an external static pd , can also lead to nonuniform concentration. For such situations the small-signal ac continuum equations cannot be solved exactly. Yet these situations are often of great experimental importance. Response has therefore been calculated for a few situations by numerical methods. In addition to work on supported situations[48], Buck and Brumleve[49] and Franceschetti and the author have applied such methods to the unsupported case[50, 51]. For a binary electrolyte, both blocking and partly blocking electrode conditions were investigated using both CJ and BV boundary conditions with charge of both signs or of only one sign mobile. Without specific adsorption effects included, it was found that the circuit of Fig. 2, with only C_∞ , R_∞ , C_R , and R_R non-zero, was able to fit the results quite well for an appreciable range of positive and negative applied dc bias. The elements of the circuit depended on the bias in ways which were consistent with the presence of accumulation and depletion layers near the electrode(s)[51]. No such circuit fitting is available for comparison in the supported situation.

COMPLEX NONLINEAR LEAST SQUARES FITTING

Fitting of IS data to a model in order to obtain model parameter estimates is crucial to the identification, interpretation, and quantification of the physical processes leading to the data. The most powerful available fitting method, which also yields goodness-of-fit and parameter standard deviation (SD) estimates, is CNLS[52-56]. Since the various CNLS approaches have been described and compared in[56], here only recent improvements to the author's CNLS program will be discussed. The latest version of this program, LEVM, may be obtained from the author's department for a nominal fee, and old versions will be updated free upon request.

Figure 3 shows some of the main fitting circuits available in LEVM. In these circuits, "DE" indicates a DCE, and "DAE" a DCE involving a general exponential (EDA) or gaussian distribution (GDA) of activation energies, or the exact small-signal blocked-electrode model[25]. All the circuits allow the input of impedance of the measuring instrument to be included when appropriate. In addition, when desired, the actual fitting can be carried out for the combination of a given circuit and a known reference impedance, as in a frequency response analyser. A very important feature of these circuits is that only those elements which are taken non-zero in the CNLS fitting appear in a given circuit. These may be defined as fixed in value, free, or free and positive only. Because of this feature, the five circuits shown

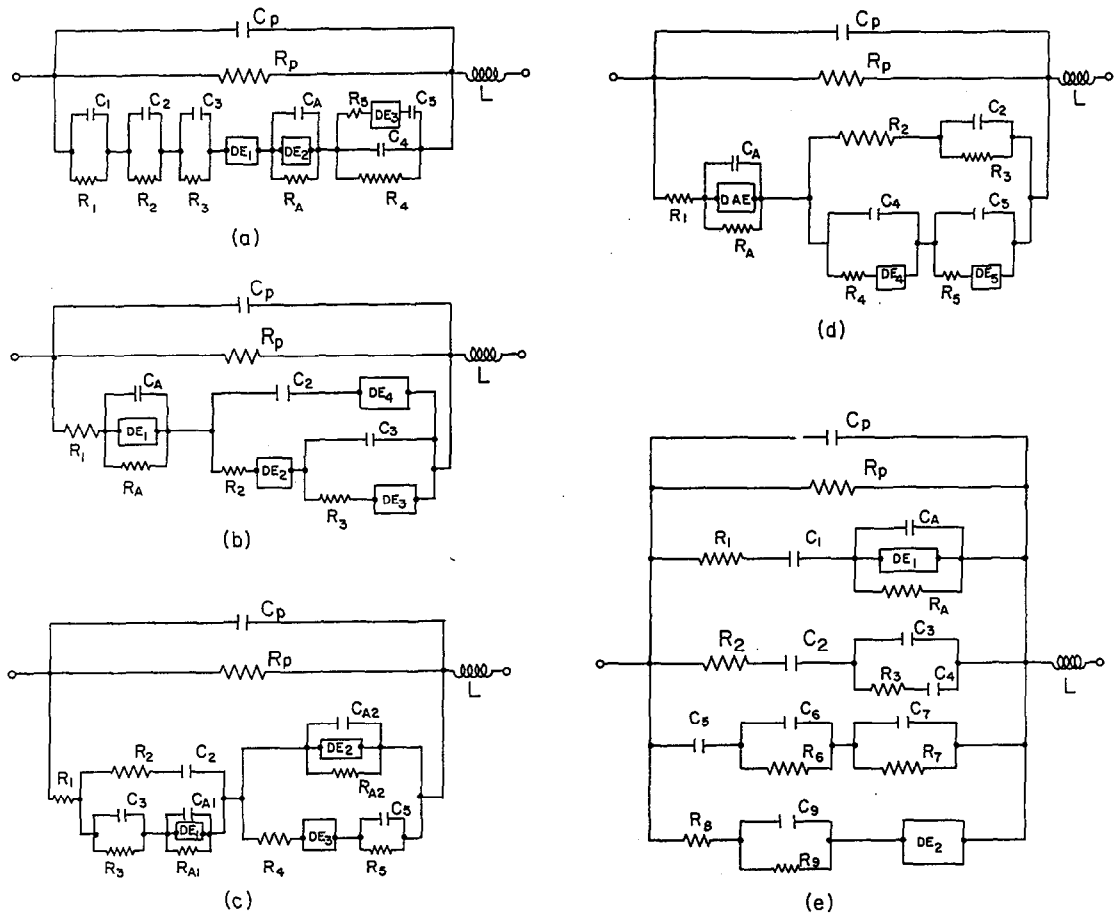


Fig. 3. Five general equivalent circuits for use in forming simpler circuits for CNLS fitting with the LEVM program. Here DE indicates one of many distributed circuit elements, and DAE denotes a full distribution of activation energies distributed element.

actually encompass tens of thousands of circuit possibilities. In addition, there are ten different DCE choices available for any DE, each involving up to six parameters[56]. The input data may be in any of the four basic immittance forms, either rectangular or polar, and may be fitted in any other form desired. The program may also be used to fit WW, EDAE, and GDAE transient response.

A new feature of LEVM is optimization for complex data fits. When invoked, it automatically adjusts the weighting of the real and imaginary parts of the data in order to make them contribute optimally to the final fit. Another important addition is the choice of robust regression fitting rather than least squares fitting. It is particularly appropriate when errors are large. Even more important are the new fixed and automatic weighting options. First consider the form of the errors likely to be present in IS data. Let $I_c(\omega)$ denote experimental IS data, $I_{fc}(\omega)$ the corresponding error-free data (generally unknown), and $I_f(\omega, \theta)$ denote a fitting model involving the parameter set θ . Further, take $I_{fc}(\omega, \theta)$ as the correct fitting model and let θ_0 represent the set of exact parameter values. Then $I_{fc}(\omega, \theta_0) = I_c(\omega)$. Now the data, models, and errors, $\epsilon_j \equiv \epsilon'_j + i\epsilon''_j$ (not the complex dielectric constant), may be related, for $j = 1, 2, \dots, N$, by:

$$I_c(\omega_j) \equiv I'_c(\omega_j) + iI''_c(\omega_j) = I_{fc}(\omega_j, \theta_0) + \epsilon_j, \quad (5)$$

which is to be fitted, using CNLS, by the model $I_f(\omega_j, \theta)$, which may or may not be the usually unknown $I_{fc}(\omega_j, \theta)$. We now make the plausible assumption that the errors may involve independent resolution and power-law components ([57], pp. 5 and 57) of the form

$$\epsilon'_j = \alpha_r P_{1j}(0, 1) + \sigma_r P_{2j}(0, 1) \times \text{sgn}\{I'_{fc}(\omega_j, \theta_0)\} [|I'_{fc}(\omega_j, \theta_0)|^{\xi_0}], \quad (6)$$

and

$$\epsilon''_j = \alpha_r P_{1j}(0, 1) + \sigma_r P_{3j}(0, 1) \times \text{sgn}\{I''_{fc}(\omega_j, \theta_0)\} [|I''_{fc}(\omega_j, \theta_0)|^{\xi_0}]. \quad (7)$$

The P 's are independent probability distributions, assumed normal here, of zero mean and unity SD, and P_j is a random member of P . Thus α_r and σ_r are the SD's of the additive resolution errors and of the power-law contributions, considered separately. For $\sigma_r = 0$ and $\alpha_r \neq 0$ one thus has additive errors, while for $\sigma_r \neq 0$, $\alpha_r = 0$, and $\xi_0 = 1$ the errors are of proportional form. For some situations, the choice $P_2 = P_3$ is appropriate.

In CNLS one minimizes the sum of the weighted real-part residuals and of the weighted imaginary-part ones using weights, w'_j and w''_j , which are inverses of the approximate error variances, v'_j and

v_j'' . Good parameter estimates usually require good variance estimates. The proper variance model for equations (5-7) is $v_j' = \alpha_r^2 + \sigma_r^2 |I_{fc}'(\omega_j, \theta_0)|^{2\xi_0}$ and $v_j'' = \alpha_r^2 + \sigma_r^2 |I_{fc}''(\omega_j, \theta_0)|^{2\xi_0}$. Even when I_{fc} is known, the true θ_0 parameter values are not known exactly since it is the main object of CNLS fitting to obtain good estimates of them. Therefore, the best one can usually do is to replace the $I_{fc}(\omega, \theta_0)$ in v_j' and v_j'' by $I_f(\omega_j, \theta)$. Further, the true values of α_r , σ_r , and ξ_0 are also virtually always unknown, so one must use estimates of them in the variance model. Since a good estimate of σ_r is unavailable until fitting convergence, when $\sigma_r \neq 0$ we use a scaled variance model, that above with ξ_0 replaced by its estimate ξ , σ_r replaced by unity, and α_r by the variance parameter U , an estimate of (α_r/σ_r) . Then as fitting progresses toward convergence, proper weighting will be approached if I_f equals or well approximates I_{fc} ; if U and ξ are good choices, and if the final set of θ values well approximates the θ_0 set. Then, for example, S_r , the SD of the overall fit, will be a good estimate of σ_r for the $\alpha_r = 0$ situation.

If the errors are believed to be additive, one would set $\sigma_r = 0$ and use any constant value of U , leading to unity weighting (UWT). On the other hand, when σ_r and ξ are non-zero, consider first two possible choices for the I_{fc} components in the variance model: components of I_f or of I_c . We shall identify the first choice, where the weighting involves function values, and thus varies with each iteration as the θ 's vary, by adding the letter "F" to the weight designation. Further, for either choice, when $|I|$, rather than the components of I , is used in both v_j' and v_j'' , yielding modulus weighting, an "M" will be added to the designation. Thus in the common and important situation where α_r is negligible and ξ is set to 1, the constant-coefficient-of-variation model, one may define and use PWT, MWT, FPWT, or FMWT. When ξ is non-zero and free and U is zero or free, we may, analogously, define the general weights, GWT, GMWT, FGWT, and FGMWT. Thus far, only UWT, PWT, or MWT have been used in most CNLS fitting, but Monte Carlo CNLS fitting with the other weightings suggests that they may often be more appropriate. All the above weighting possibilities are incorporated in the new version of the CNLS program.

Modulus-type weighting is primarily appropriate when $|I_c'|$ and $|I_c''|$ remain comparable in size for the full data set, often not the case. Although the UWT, PWT, and MWT weights involve I_c and so do not vary during CNLS iteration, fitting with varying weights is essentially no more difficult than with constant ones. More important, PWT and MWT generally lead to much more bias in parameter estimates than do FPWT and FMWT, even when $I_f = I_{fc}$ (see below). Only when one is quite uncertain about how well the fitting model I_f approximates I_{fc} is PWT more reasonable than FPWT. Thus it is often appropriate to use PWT in preliminary fitting; then if a seemingly good fit is obtained, FPWT should be used for final fitting.

Monte Carlo (MC) simulation CNLS fitting of the circuits of Fig. 1 has been carried out to quantify the above bias differences. Exact values of the parameters (in ohms and farads) of these circuits are listed in

Table 1. Those for the ambiguous circuits D, A, C, and J have been selected to yield exactly the same response for all the $N = 33$ frequencies used, extending over the range $10^{-4} \leq \omega \leq 10^4$, for each of them. The impedance of the ZC DCE in Fig. 1 is $R_2/[1 + (i\omega\tau)^\psi]$. Each individual k out of the total K simulation Z-level fits involves data with normal random errors satisfying equations (6) and (7), and, initially, with the fixed values $\alpha_r = 0$, $U = 0$, and $\xi_0 = \xi = 1$. Let p denote the total number of free parameters, $m = 1, 2, \dots, p$, and define the relative error of the m^{th} parameter in the k^{th} fit, e_{mk} , the estimate relative bias of the m^{th} parameter, b_m , and the MC estimated relative SD of the m^{th} parameter, s_{am} , as

$$e_{mk} \equiv [\theta_{mk} - \theta_{0mk}]/\theta_{0mk}, \tag{8}$$

$$b_m \equiv K^{-1} \sum_{k=1}^K e_{mk}, \tag{9}$$

$$s_{am} \equiv (f/K) \sum_{k=1}^K |e_{mk}|, \tag{10}$$

where $f \simeq 1.483$ for a normal distribution and $2/\sqrt{3}$ for a uniform one. The first value has been used for the results below. It is, however, appropriate only for negligible bias and depends somewhat on bias when this is not negligible. Simulation also yields a linearized estimate of parameter relative SD's s_{rm} , which may not depend so directly on the parameter error distribution[56]. Now although b_m/σ_r depends strongly on σ_r for the model parameters, the ratios $r_{am} \equiv s_{am}/\sigma_r$ and $r_{rm} \equiv s_{rm}/\sigma_r$ are independent of it for the D, A, C, and J circuits. To obtain well-defined estimates of s_{am} or s_{rm} one needs a K of $O(10^4)$, while good b_m estimates often require a K of $O(10^6)$ when $|b_m|$ is $\ll 1$.

The r values listed in Table 1 for each θ_m value provide a measure of how well a given parameter is likely to be determined in a single CNLS fit: $\theta_m[1 \pm S_r r_m]$. They are thus concerned with potential precision of estimates, while a bias is a measure of the limits of ultimate accuracy of estimation (unless it is known and can be taken account of). Note that

Table 1. Exact (0 to 2 decimal places) and approximate parameter values and corresponding values of the r ratios for FPWT fitting. τ and ψ apply only for E

| θ_m r_m | Circuits | | | | |
|---------------------|----------|-------|--------|-------|-----------------|
| | D | J | C | A | E |
| R_1 | 0.3 | 1 | 1 | 0.530 | 10^6 |
| r_a | 0.47 | 0.13 | 0.13 | 0.22 | 0.20 |
| r_r | 0.43 | 0.16 | 0.16 | 0.22 | 0.21 |
| C_1 | 1 | 1 | 1.49 | 2.530 | 10^{-12} |
| r_a | 0.15 | 0.15 | 0.35 | 0.49 | 0.23 |
| r_r | 0.17 | 0.17 | 0.32 | 0.44 | 0.25 |
| R_2 | 0.7 | 0.429 | 0.0463 | 0.947 | 2×10^6 |
| r_a | 0.26 | 0.67 | 1.62 | 1.66 | 0.20 |
| r_r | 0.25 | 0.59 | 1.39 | 1.42 | 0.22 |
| $C_2, (\tau)$ | 1 | 0.49 | 3.041 | 1.654 | 1 |
| r_a | 1.17 | 1.03 | 0.75 | 0.85 | 1.42 |
| r_r | 1.00 | 0.88 | 0.64 | 0.73 | 1.20 |
| ψ | — | — | — | — | 0.3 |
| r_a | — | — | — | — | 0.29 |
| r_r | — | — | — | — | 0.25 |

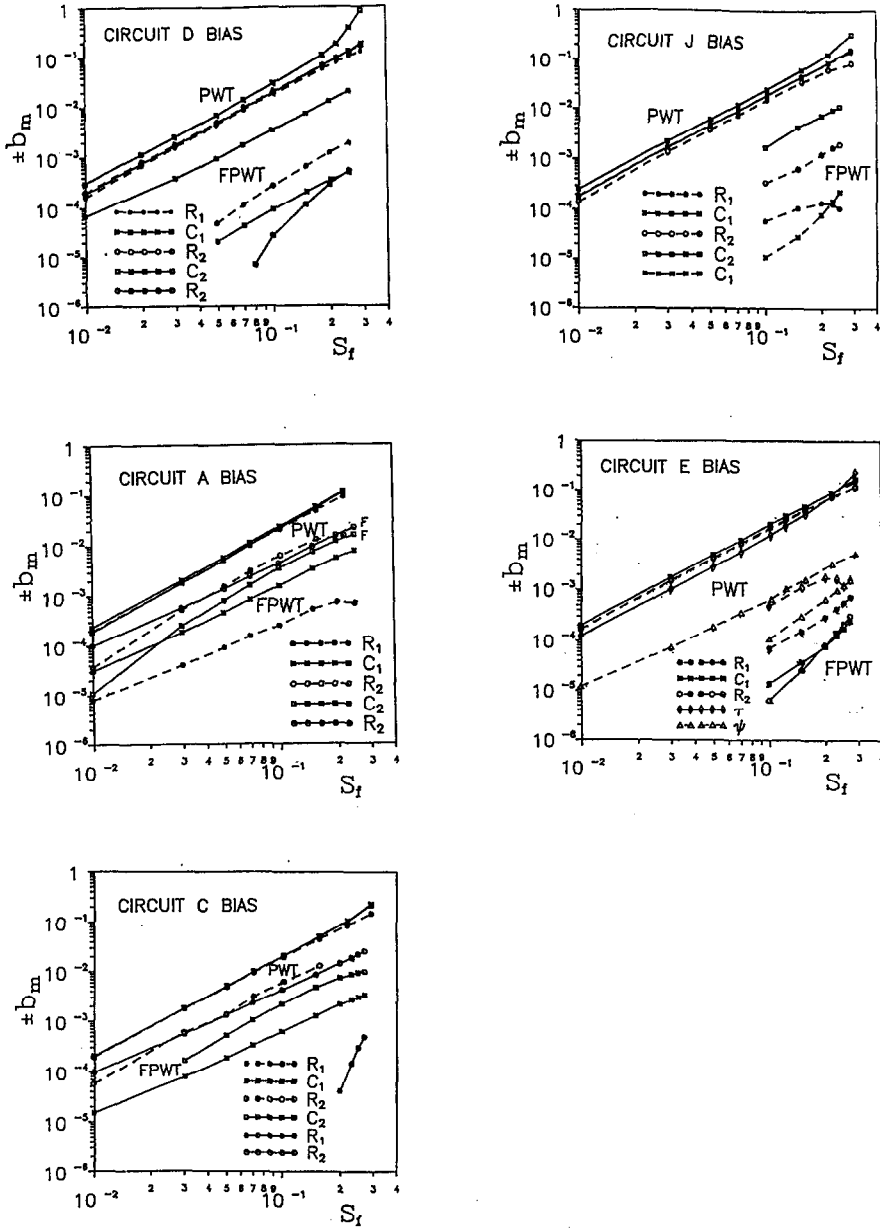


Fig. 4. Dependence of the bias b , estimated from simulation fitting on the mean standard deviation of K fits s_f , for the five circuits of Fig. 1 and for PWT and FPWT. Dotted lines indicate $b < 0$.

parameters which are equal in value and play the same role in two circuits, such as $J-R_1$ and $C-R_1$, necessarily have equal r values. It is found that the r values for FMWT are appreciably larger than those for FPWT (eg, for D- R_1 , the FMWT values of r_a and r_r are 1.93 and 1.49, respectively), and those of UWT are very much larger. Thus in a fit of a single data set, FMWT parameter estimates will actually be appreciably more uncertain than FPWT ones for data of the present type.

Bias results vs S_f for the five circuits are shown in Fig. 4 for the variance model parameter values given above. There are three possible sources of bias: wrong fitting model, wrong variance model, and the intrinsic nonlinearity of the fitting model. The first source is absent here since we take $I_f(\omega, \theta) = I_c(\omega, \theta)$. The

figure shows that wrong weighting, here PWT, usually leads to bias 10–100 times or more larger than the residual nonlinearity bias of the correct FPWT. The results of Fig. 4 and Table 1 show that for $\sigma_f \geq 0.1$, for the D or J circuit the PWT C_1 bias even exceeds its statistical uncertainty, s_a or s_r . Further, UWT yields appreciably greater bias even than PWT. Although the large PWT bias results are nearly the same for the four ambiguous circuits, appreciable differences appear for FPWT. Note also that the FPWT bias results are so small that generally $(b_m \pm S_f r_m) \approx S_f r_m$ for even large values of S_f , and such bias is thus negligible compared to normal statistical uncertainty.

All previous CNLS work has involved constant values of U and ξ , most often with either UWT or

PWT. For real situations, however, appropriate values of U and ξ are usually unknown. Therefore, it is desirable to allow them to be free parameters of the least squares fit as well as the θ 's. By modifying and extending an approach discussed in Refs.[57, 58], it has been possible to do so, and the results, which usually yield U and ξ estimates with high resolution, are incorporated in LEVM. Most appropriately, the U and ξ estimates from such a fit may then be used as fixed values in a final fitting. For a situation where $\alpha_r = 0$, $\xi_0 = 1$, and ξ is free, one finds for the four equivalent circuits of Fig. 4 that FPWT yields $b_\xi \approx 2.7 \times 10^{-3}$, essentially independent of σ_r , $r_{a\xi} \approx 2.63$, and $r_{r\xi} \approx 1.52$, both quite small enough that ξ estimates will be useful in most cases of interest. Results for the E circuit are comparable. These relative results are all independent of the data scaling as expected. Present space limitations, prohibit a detailed discussion of the important situation where both α_r and σ_r are nonzero and fitting involves both U and ξ free. MC results show, for example, that for the D, A, C, and J circuits $U \approx \alpha_r/\sigma_r$, as expected, and $s_{rU} \approx 0.16$, independent of σ_r and α_r values. One finds that for these four circuits with $\alpha_r = 10^{-3}$, $\xi_0 = 1$, and $\sigma_r = 0.1$, then $U \approx 0.01$, $b_\xi \approx 0.03$, $s_{a\xi} \approx 0.08$, and $s_{r\xi} \approx 0.07$. No convergence problems appear for any of the circuits until α_r becomes comparable with σ_r ($|I_e|_{\max}$)⁵⁰. Then, ordinary UWT is appropriate.

Acknowledgements—I much appreciate the preliminary Monte Carlo work of Larry D. Potter, Jr. The support of the Army Research Office is gratefully acknowledged.

REFERENCES

1. M. Sluyters-Rehbach and J. H. Sluyters, in *Electroanalytical Chemistry*, Vol. 4, (Edited by A. J. Bard), pp. 1–127. Marcel Dekker, New York (1970).
2. R. Parsons, *Adv. Electrochem. electrochem. Engng* **7**, 177 (1970).
3. J. R. Macdonald, in *Superionic Conductors*, (Edited by G. D. Mahan and W. L. Roth), pp. 81–97, Plenum Press, New York (1976).
4. J. R. Macdonald, in *Electrode Processes in Solid State Ionics*, (Edited by M. Kleitz and J. Dupuy), pp. 149–183, Reidel, Dordrecht, Holland (1976).
5. R. D. Armstrong, M. F. Bell and A. A. Metcalfe, *Electrochem. chem. Soc. Spec. Rep.* **6**, 98 (1978).
6. W. I. Archer and R. D. Armstrong, *Electrochem. Chem. Soc. Spec. Rep.* **7**, 157 (1980).
7. J. R. Macdonald, *IEEE Trans. Electrical Insulation EI-15*, 65 (1981).
8. F. Mansfield, *Corrosion* **37**, 301 (1981).
9. F. Mansfield, M. W. Kendig and S. Tsai, *Corrosion* **38**, 570 (1982).
10. R. P. Buck, *Ion-Selective Electrode Rev.* **4**, 3 (1982).
11. M. Sluyters-Rehbach and J. H. Sluyters, in *Comprehensive Treatise of Electrochemistry*, Vol. 9, (Edited by E. Yeager, J. O'M. Bockris, B. E. Conway and S. Sarangapani), pp. 177–292. Plenum Press, New York (1984).
12. J. R. Macdonald (Editor), *Impedance Spectroscopy—Emphasizing Solid Materials and Systems*. Wiley-Interscience, New York (1987)*.
13. J. R. Macdonald, *J. electroanal. Chem.* **223**, 25 (1987).
14. H. Nyquist, *Bell Syst. tech. J.* **11**, 126 (1932).
15. G. Jaffé, *Phys. Rev.* **85**, 354 (1952).
16. H. Chang and G. Jaffé, *J. chem. Phys.* **20**, 1071 (1952).
17. J. R. Macdonald, *Phys. Rev.* **92**, 4 (1953).
18. R. J. Friauf, *J. chem. Phys.* **22**, 1329 (1954).
19. R. P. Buck, *J. electroanal. Chem.* **210**, 1 (1986).
20. J. H. Beaumont and P. W. M. Jacobs, *J. Phys. Chem. Solids* **28**, 657 (1967).
21. J. R. Macdonald, *Trans. Faraday Soc.* **66**, 943 (1970).
22. J. R. Macdonald, *J. chem. Phys.* **54**, 2026 (1971)†.
23. J. R. Macdonald, *J. electroanal. Chem.* **32**, 317 (1971); **47**, 182 (1973); **53**, 1 (1974).
24. J. R. Macdonald, *J. chem. Phys.* **58**, 4982 (1973); **61**, 3977 (1974).
25. J. R. Macdonald, *J. electrochem. Soc.* **135**, 2274 (1988).
26. D. R. Franceschetti and J. R. Macdonald, *J. electroanal. Chem.* **82**, 271 (1977).
27. D. R. Franceschetti and J. R. Macdonald, *J. electroanal. Chem.* **87**, 419 (1978)‡.
28. J. R. Macdonald and D. R. Franceschetti, *J. electroanal. Chem.* **99**, 283 (1979).
29. J. R. Macdonald, *J. appl. Phys.* **62**, R51 (1987).
30. R. L. Hurt and J. R. Macdonald, *Solid St. Ionics* **20**, 111 (1986).
31. J. R. Macdonald, *Bull. Am. phys. Soc.* **30**, 587 (1985).
32. J. R. Macdonald, *J. appl. Phys.* **58**, 1955 (1985).
33. J. R. Macdonald, *J. appl. Phys.* **58**, 1971 (1985)§.
34. J. R. Macdonald, *J. appl. Phys.* **61**, 700 (1987).
35. J. R. Macdonald, *J. appl. Phys.* **65**, 4845 (1989).
36. A. Le Mehaute and G. Crepy, *Solid St. Ionics* **9/10**, 17 (1983).
37. S. H. Liu, *Phys. Rev. Lett.* **55**, 529 (1985).
38. J. C. Wang, in Sixth International Conference on Solid State Ionics, Garmisch-Partenkirchen (1987); *Solid St. Ionics* **28/30**, 1436 (1988).
39. B. Sapoval, J.-N. Chazalviel and J. Peyriere, *Phys. Rev. A* **38**, 5867 (1988).
40. J. B. Bates, Y. T. Chu and W. T. Stribling, *Phys. Rev. Lett.* **60**, 627 (1988).
41. J. R. Macdonald and D. R. Franceschetti, *J. chem. Phys.* **68**, 1614 (1978).
42. D. R. Franceschetti and J. R. Macdonald, *J. electroanal. Chem.* **101**, 307 (1979).
43. D. R. Franceschetti and J. R. Macdonald, *J. electrochem. Soc.* **129**, 1754 (1982).
44. J. R. Macdonald and C. A. Hull, *J. electroanal. Chem.* **165**, 9 (1984).
45. R. Labes and H. Lullies, *Pflugers Arch. ges. Physiol.* **230**, 738 (1932).
46. H. Lullies, *Biol. Rev.* **12**, 338 (1937).
47. K. S. Cole, *Membranes, Ions, and Impulses*, pp. 167–203. University of California Press, Berkeley, CA (1972).
48. S. W. Feldberg, *J. electroanal. Chem.* **3**, 199 (1969).
49. T. R. Brumleve, *J. electroanal. Chem.* **90**, 1 (1978).
50. D. R. Franceschetti and J. R. Macdonald, *J. electroanal. Chem.* **100**, 583 (1979).
51. D. R. Franceschetti and J. R. Macdonald, *Proceedings of the 3rd Symposium on Electrode Processes* (Edited by S. Bruckenstein, J. D. E. McIntyre, B. Miller and E. Yeager). *Proc. electrochem. Soc.* **80-3**, 94–114 (1980).
52. J. R. Macdonald and J. A. Garber, *J. electrochem. Soc.* **124**, 1022 (1977).
53. J. R. Macdonald, J. Schoonman and A. P. Lehnen, *J. electroanal. Chem.* **131**, 77 (1982).

*The l in equations (33), (35) and (36) on pp. 105–106 should be replaced by $l/2$.

†Errata published in *J. chem. Phys.* **56**, 681 (1972). In addition the words "intrinsic" and "extrinsic" are improperly used in this work (in place of "intensive" and "extensive").

‡The term $(1 - i - n)$ in equation (8) of this paper should be $(1 + i - n)$.

§The $\exp(-N_{ii} E)$ term in equation (17) should be replaced by $\exp(-\eta_{ii} E)$ and the "+" sign in equation (24) should be "=".

54. Y.-T. Tsai and D. H. Whitmore, *Solid St. Ionics* **7**, 129 (1982).
55. B. A. Boukamp, *Solid St. Ionics* **20**, 31 (1986).
56. J. R. Macdonald and L. D. Potter Jr, *Solid St. Ionics* **23**, 61 (1987).
57. R. J. Carroll and D. Ruppert, *Transformation and Weighting in Regression*. Chapman and Hall, New York (1988).
58. S. L. Beal and L. B. Sheiner, *Technometrics* **30**, 327 (1988).