PRACTICAL TRANSFER FUNCTION ESTIMATION
AND ITS APPLICATION TO
WIDE FREQUENCY RANGE REPRESENTATION OF TRANSFORMERS

A. Oguz Soysal*   Adam Semlyen

Department of Electrical Engineering
University of Toronto
Toronto, Ontario, Canada, M5S 1A4

* On leave from Karadeniz Technical University, Trabzon, Turkey

Abstract - A widely applicable, general methodology for estimation of transfer function parameters from frequency response data is presented. The procedure is based on the solution of a linear least squares problem by the singular value decomposition (SVD). The condition of the problem is discussed and approaches referred to as shifting and scaling are introduced to reduce the condition number. To extend the application to practical cases with measurement errors and/or a large number of poles, a partitioned estimation method with Gauss-Seidel iterations is developed. Finally, an iterative improvement process with constraints on the poles is applied to increase the accuracy and to avoid the possibility of obtaining unstable poles. The application of the suggested method of estimation to the representation of transformers is presented with practical examples. Either transfer function or state equation representation can be obtained for transformers described by their terminal frequency responses.

Keywords: System identification, Parameter estimation, Least squares approximation, Frequency response, Transformer modeling.

INTRODUCTION

Accurate representation of physical systems by a transfer function is often needed in various fields of engineering for purposes of analysis, design, or simulation. In many cases where sufficient information about the structure of the system is not available, frequency response measurements can be used as a convenient data base for estimation of the transfer function parameters. System identification by terminal characteristics may also be useful for the lumped parameter modeling of distributed systems, reduction of model order, and simplification of complex systems. In this paper, different methodologies for transfer function estimation are discussed. Novel procedures are proposed and examined, some superior to existing ones, developed for diverse engineering applications.

In the power engineering area, wide frequency range modeling of transformers and reactors by frequency domain external measurements is sometimes required for the study of electromagnetic transients. Besides, the "black-box" representation of such equipment is particularly important in insulation coordination applications involving HV and EHV systems, where precise prediction of possible system overvoltages is essential to achieve an economical and reliable design.

In the study of electromagnetic transients, the components must be represented in a wide frequency range. This introduces numerical difficulties with most of the available methods. The estimation process proposed in this paper is particularly appropriate for identification of transfer function parameters from wide frequency band observation data, obtained by terminal measurements of power system components such as transformers, reactors, rotating machines, and transmission lines. However, due to its improved accuracy and the achieved stability, the proposed method can also be applied with advantage to a large class of estimation problems in various engineering fields.

Background

Considerable effort has been devoted since the 1950s to the development of methods for transfer function synthesis from frequency response observations. The frequency domain identification problem is based on the estimation of a rational complex function, with real coefficients, to fit a given set of complex data. The nonlinear nature of the problem has yielded different formulations and solution methods.

Levy [1] suggested a linearization procedure and developed a complex curve fitting method. The proposed method was based on the *least squares approach* and the use of partial derivatives to minimize a quadratic error function yielded a formulation equivalent to the use of *normal equations*. In Levy's method, the error function is inherently weighted due to the multiplication of the equations by the denominator and the resulting estimation is biased. Besides, because of the normal equation type formulation, the problem was poorly conditioned. Sanathanan and Koerner [2] improved Levy's method by introducing an iterative technique to remove the undesired weighting. Presented results show clearly that reverse weighting reduces the errors in the fitting. However, Payne [3] reported that in some cases, both the Levy and the Sanathanan-Koerner methods may give rise to right-half-plane poles for systems known to be stable. To avoid this disadvantage, Payne suggested to consider some time domain properties of the system, such as steady state errors, in conjunction with the frequency domain data.

Lawrence and Rogers [4] developed a sequential algorithm based on Levy's linearization procedure. The proposed algorithm allows point-by-point determination of the transfer function parameters and, in the case of addition of new observation data, the previously identified parameters are not reevaluated but updated. Another major improvement of this method is that a matrix inversion or solution of a set of equations is not needed. However, in the case of higher order transfer functions, a large number of observations is needed to obtain an accurate solution, otherwise iterations become necessary. Stahl [5] proposed two different procedures, namely *matrix adaptation* and *direct solution* methods. The Matrix adaptation method is an iterative process based on Levy's error function. The coefficient matrix elements are functions of parameters to be determined and are updated at each iteration. The direct solution uses a logarithmic cost function to be minimized by the simplex method.

Whitfield [6] presented the previous approaches in a unified structure and discussed them in terms of convergence and error considerations. Whitfield developed the *integral approach* method where a time domain integral error criterion is discretized by the trapezoidal rule. The final formulation is in the form of an overdetermined set of linear equations which are solved in the least squares sense by Householder decomposition. The non-iterative integral approach has the tendency to give more emphasis to the low frequency observations particularly in the presence of measurement errors. Comparative results of Whitfield show that the methods proposed by Levy, Sanathanan-Koerner, and Stahl may yield unstable transfer functions for stable systems and the iterative methods may not converge in some cases.

In this work, the basic formulation of Levy is modified to obtain an overdetermined set of linear equations. The limitations of finite digit computation due to the ill-conditioned nature of the equations is discussed and it is shown that the numerical condition can be improved by appropriate *scaling. Singular Value Decomposition (SVD)*, rather than normal equations, is used for direct solution of the overdetermined equations in the least squares sense. An appropriate *partitioning process* is presented for sequential identification of poles or groups of poles that are widely separated over the frequency range of the observations. Finally, a further *iterative improvement process* is introduced to reduce the calculation errors and to avoid the possibility of obtaining unstable poles. By these procedures, the weaknesses of the existing methods, discussed above, have been overcome, as demonstrated in the examples and applications presented in the paper.

BASIC THEORY

Least Squares Formulation

The transfer function of a linear time invariant single-input, single-output system may be written as :

$$H(s) = \frac{a_0 + a_1 s + a_2 s^2 + \cdots + a_\mu s^\mu}{b_0 + b_1 s + b_2 s^2 + \cdots + b_\nu s^\nu} \qquad (1)$$

For a non-integrating system we can assume $b_0 = 1$ without loss of generality. If $H(s)$ is strictly proper (i.e. $\lim H(s) = 0$ for $s \to \infty$), then $\mu < \nu$, normally, $\mu = \nu - 1$. Besides, let us assume that $m/2$ complex observation values are available as $h_k = p_k + jq_k$ for the observation frequencies ω_k. By substituting $s = j\omega_k$ in (1) we obtain

$$\frac{N_r(\omega_k) + jN_i(\omega_k)}{1 + D_r(\omega_k) + jD_i(\omega_k)} \approx p_k + jq_k \qquad (2)$$

$$N_r(\omega_k) = a_0 - a_2\omega_k^2 + a_4\omega_k^4 - \cdots$$
$$N_i(\omega_k) = a_1\omega_k - a_3\omega_k^3 + a_5\omega_k^5 - \cdots$$
$$D_r(\omega_k) = -b_2\omega_k^2 + b_4\omega_k^4 - b_6\omega_k^6 + \cdots \qquad (3)$$
$$D_i(\omega_k) = b_1\omega_k - b_3\omega_k^3 + b_5\omega_k^5 - \cdots$$

Multiplying both sides of (2) with the denominator and equating real and imaginary parts, we obtain

$$N_r(\omega_k) - p_k D_r(\omega_k) + q_k D_i(wk) \approx p_k \qquad (4)$$

$$N_i(\omega_k) - q_k D_r(\omega_k) - p_k D_i(\omega_k) \approx q_k \qquad (5)$$

Equations (4) and (5) are linear in the $n = 2\nu$ parameters a_i ($i=0,...,\nu-1$) and b_i ($i=1,...,\nu$). For the m real observations we can write the resulting set of m equations in n unknowns, as

$$\mathbf{Ax} \approx \mathbf{b} \qquad (6)$$

where

$$\mathbf{x}^T = \begin{bmatrix} a_0 & a_1 & \cdots & a_{\nu-1} & b_1 & b_2 & \cdots & b_\nu \end{bmatrix}$$

$$\mathbf{b}^T = \begin{bmatrix} p_1 & q_1 & p_2 & q_2 & \cdots & p_{m/2} & q_{m/2} \end{bmatrix}$$

$$\mathbf{A} = \text{col}\{\mathbf{A}_k\} \qquad (k = 1,...,m/2) \qquad (7)$$

$$\mathbf{A}_k = \begin{bmatrix} 1 & 0 & -\omega_k^2 & 0 & \omega_k^4 & \dots & q_k\omega_k & p_k\omega_k^2 & -q_k\omega_k^3 & -p_k\omega_k^4 & \dots \\ 0 & \omega_k & 0 & -\omega_k^3 & 0 & \dots & -p_k\omega_k & q_k\omega_k^2 & p_k\omega_k^3 & -q_k\omega_k^4 & \dots \end{bmatrix}$$

If the number m of observations is larger than the number n of parameters, (6) is an overdetermined set of linear equations. This set, which in general is of full rank, has a unique solution in the least squares sense, $\mathbf{x}^*$, that minimizes the residual $\|\mathbf{Ax} - \mathbf{b}\|$ [7]. Different methods are available for the solution of the least squares problem. An early method consists in solving the normal equations

$$\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b} \qquad (8)$$

The formulation of Levy [1], Sanathanan-Koerner [2] and Stahl [5] are equivalent to (8). Despite its simplicity, the solution of the normal equations is of reduced numerical stability and small perturbations of A and b may result in large errors in the solution. Lawson and Hanson [7] present a number of more powerful methods based on orthogonal decomposition of A. Singular Value Decomposition (SVD), which is one of the most efficient and stable techniques, will be used in this paper for the direct solution of the least squares equations. It has the further advantage that it gives directly the condition number to be discussed below.

Perturbation Analysis and the Condition of the Problem

Suppose that in (6) both A and b are subject to small perturbations ΔA and Δb respectively, which result in a deviation Δx in the solution vector x. Substitution of the perturbed values in (6) gives

$$(\mathbf{A} + \Delta\mathbf{A})(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{b} + \Delta\mathbf{b} \qquad (9)$$

If the small deviations of second order are neglected, Δx can be expressed explicitly by using the pseudoinverse $\mathbf{A}^+$ of A

$$\Delta\mathbf{x} = \mathbf{A}^+ [\Delta\mathbf{b} - (\Delta\mathbf{A})\mathbf{x}] \qquad (10)$$

By denoting with $\|\cdot\|$ any appropriate norm, we can write

$$\|\Delta\mathbf{x}\| \leq \|\mathbf{A}^+\| \, \|\Delta\mathbf{A}\| \, \|\mathbf{x}\| + \|\mathbf{A}^+\| \, \|\Delta\mathbf{b}\| \qquad (11)$$

and

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}^+\| \, \|\Delta\mathbf{A}\| + \frac{\|\mathbf{A}^+\| \, \|\Delta\mathbf{b}\|}{\|\mathbf{x}\|} \qquad (12)$$

Let us define

$$N_c = \|\mathbf{A}\| \, \|\mathbf{A}^+\| \qquad (13)$$

The relative error in the solution can be expressed in terms of N_c and the relative deviations of parameters as

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq N_c \left(\frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} \right) \qquad (14)$$

N_c is the *condition number of* A and can be considered as a measure of the amplification of perturbations on A and b throughout the computation of x, independently of the chosen procedure. The perturbation on the parameters may arise either due to the round-off errors of the finite digit computation or the inevitable noise in the measurements. If the condition number is too large, the problem is said to be ill-conditioned and the results are generally unreliable. Besides, it should be noted that in the normal equations approach (8) the condition number is that of $\mathbf{A}^T\mathbf{A}$, the square of the condition number of A proper in (6).

Although the norm definition is general in (11)-(14), it is common to use the euclidian norm for vectors. On the other hand, the condition number may be calculated in terms of the singular values of A

$$N_c = \frac{\sigma_{max}}{\sigma_{min}} \qquad (15)$$

where σ_{max} and σ_{min} denote the biggest and smallest singular values, respectively [8]. Equation (15) allows the direct calculation of the condition number without computing the pseudoinverse, if SVD is used to solve (6).

A close examination of the structure of the A matrix given in (7) will show that the condition number depends basically on the following factors:

– Order of the transfer function

– Frequency range of the observations

– Observation data (p_k, q_k)

Condition numbers for some typical example models will be given in Table 1. It can be verified that the condition number increases as ω_{max}^ν, where ω_{max} is the maximum angular frequency of observations and ν the degree of the transfer function denominator. As in the examples presented here, in many practical problems the condition number may be in the order of 10^{16} or higher which makes an accurate solution impossible even with double precision computation. As a consequence, it is obvious that the condition of the problem must be improved before applying any numerical method for solution.

The ill-conditioning of A in (6) comes from the nature of the problem formulation. The iterative reverse weighting process of Sanathanan-Koerner [2] and matrix adaptation algorithm of Stahl [5] helped to improve the condition of the problem to some extent, but the use of the normal equations had already increased the condition number to the square of its original value!

In order to improve the condition of the problem, (6) may be preprocessed by scaling or shifting as described below.

Scaling

In general terms, scaling consists in left or right multiplication of the matrix A with an appropriate diagonal matrix D, an approach that can be called *row* or *column* scaling, respectively.

Van der Sluis [9],[10] investigated the stability of solutions of linear algebraic systems and presented methods for equilibration of rectangular matrices. It has been proved in [9] that the condition number of a rectangular matrix can be approximately minimized by either row or column scaling. Efficient scaling results in all rows or all columns having equal euclidian norms, respectively. It is, however, clear that despite the advantage of improving the condition of A, row scaling has the side effect of introducing an undesired weighting of the equations. Besides, it turns out that row scaling is not as efficient as column scaling. For the latter, expression (6) can be written as

$$\mathbf{ADD}^{-1}\mathbf{x} = \mathbf{b} \qquad (16)$$

With $\hat{\mathbf{A}} = \mathbf{AD}$ and $\hat{\mathbf{x}} = \mathbf{D}^{-1}\mathbf{x}$ we obtain,

$$\hat{\mathbf{A}}\hat{\mathbf{x}} = \mathbf{b} \qquad (17)$$

Similarly to (6), (17) also has a unique solution in the least squares sense, $\hat{\mathbf{x}}^*$, which minimizes $\|\hat{\mathbf{A}}\hat{\mathbf{x}} - \mathbf{b}\|$. Note that the change of variables has not changed the residual. After solving (17), the transfer function parameters can be simply calculated by using $\mathbf{x} = \mathbf{D}\hat{\mathbf{x}}$.

Shifting

In the cases where the median of the observations is far from the origin, compared to the bandwidth, the large values of ω may result in a considerable increase in the condition number of A. Experiments have shown that a change of variables, by shifting of the origin approximately to the median $j\omega$ value, can

improve the condition of the problem in such cases. Let us define a new complex variable $\bar{s}$ resulting from a shift of the origin to the point p_0 in the s-plane:

$$\bar{s} = s - p_0 \qquad (18)$$

Substitution of s defined in terms of the new origin into (1) gives

$$H(\bar{s}) = \frac{\bar{a}_0 + \bar{a}_1\bar{s} + \bar{a}_2\bar{s}^2 + \cdots + \bar{a}_\mu\bar{s}^\mu}{\bar{b}_0 + \bar{b}_1\bar{s} + \bar{b}_2\bar{s}^2 + \cdots + \bar{b}_\nu\bar{s}^\nu} \qquad (19)$$

Here, all parameters $\bar{a}_i$ and $\bar{b}_i$ are complex. Once these parameters are estimated by the least squares approach as described in the previous section, the original a_i and b_i parameters can be calculated by solving the linear equations

$$\mathbf{Px} = \bar{\mathbf{x}} \qquad (20)$$

where, $\mathbf{P} = \text{diag}\{\mathbf{P}_a, \mathbf{P}_b\}$. The matrices $\mathbf{P}_a$ and $\mathbf{P}_b$ are upper triangular with elements similar to those in the well known *Pascal triangle*. $\bar{\mathbf{x}}$ is similar to $\mathbf{x}$ defined in (7). The triangular form makes it possible to obtain a simple and fast solution.

Results

The improvement achieved by row scaling, column scaling and shifting were checked by means of a number of examples. The numerical results have shown that:

- The improvement achieved by shifting is not significant in most cases. However, when the observation interval is far from the origin, it is helpful to shift the origin approximately to the center of the observations. In some of the examples considered, where the original condition numbers were in the order of 10^{20}, the improvement was in the order of 10-10^3.

- Row scaling has improved the condition number in excessively damped systems where the ill-conditioning is mainly related to the nature of the frequency response rather than the frequency range. In slightly damped examples with wide frequency range measurements, which is the case for most power system equipment, the improvement of the row scaling was not sufficient to reduce the condition number below 10^{16}, which can be considered as an extreme for double precision computations.

- Column scaling was very efficient in most of the examples, particularly when the ill-conditioning was the result of the wide frequency range. In many examples with excessively high condition numbers, as shown in Table 1, significant improvements could be obtained by column scaling which resulted in condition numbers below 10^7.

The above mentioned examples have shown that column scaling is the most efficient method to improve the condition of the A matrix of equation (6). Besides, as it is seen from (16), the column scaling does not introduce any additional weighting to the linearized equations (6).

Direct Solution by Singular Value Decomposition

Matrix decomposition methods are known to be more powerful and efficient for the solution of linear least squares problems than the use of normal equations [7]. In this paper, the *Singular Value Decomposition* is used for direct solution in the least squares sense of the overdetermined set of equations given in (17). The rectangular $m \times n$ matrix $\hat{A}$ can be decomposed as

$$\hat{A} = USV^T \qquad (21)$$

where both U and V are orthogonal square matrices, of dimensions $m \times m$ and $n \times n$ respectively. If the rank of A is n, then

$$S = \begin{bmatrix} S_1 \\ 0 \end{bmatrix} \qquad S_1 = \text{diag}\{\sigma_i\} \quad (i = 1,...,n) \qquad (22)$$

The real nonnegative entries σ_i are the singular values of $\hat{A}$. Recalling that the euclidian length of a vector remains unchanged under multiplication with an orthogonal matrix, we can write

$$\|\hat{A}\hat{x} - b\|^2 = \|USV^T\hat{x} - b\|^2 = \|SV^T\hat{x} - U^Tb\|^2 \qquad (23)$$

Let us define

$$y = V^T\hat{x} \qquad g = U^Tb = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \qquad (24)$$

where y and g_1 are n vectors. Substituting in (23), we obtain

$$\|\hat{A}\hat{x} - b\|^2 = \|S_1y - g_1\|^2 + \|g_2\|^2 \qquad (25)$$

It is clear that the residual is minimum when y has the value

$$y^* = S_1^{-1}g_1 \qquad (26)$$

General purpose programs for SVD decomposition are available in most subroutine packages. The direct solution of (17) as described above is a straightforward process and gives satisfactory results in cases where the condition number can be reduced to such a low level that the amplification of the input errors is negligible. If measurement errors are big in the given frequency response data, or the condition number is large despite pre-conditioning, as shown in the previous section, even relatively small condition numbers may result in large errors in the estimated parameters, which lead to unrealistic values. The approaches of *partitioning* and *iterative improvement*, to be described next, can then be applied to achieve the estimation with acceptable accuracy.

Partitioning of the Frequency Scale

In many physical systems, the frequency response curves may present several minima and maxima which correspond to resonances related to discrete poles or groups of poles (see Figures 1-4). In particular, the wide frequency range response of power system components such as transformers, reactors, rotating machines, and transmission lines has in general this property. These components are in reality distributed parameter systems, which means that theoretically an infinite number of poles would be required for obtaining the measured response curve. In lumped parameter modeling, a large number of poles is normally needed to achieve acceptable accuracy. However, higher model order with a wide frequency range results in very high condition numbers. While column scaling improves the condition of the problem significantly, the error amplification will still be too large and direct fitting will not be successful.

Let us partition the frequency scale, as shown in Fig. 3, into r sections limited by frequency values corresponding to the minima of the magnitude curve. On the other hand, let us write equation (1) as a sum of r ratios of polynomials, equal to the number of partitioned sections:

$$H(s) = \sum_{i=1}^{r} H_i(s) \qquad (27)$$

By substituting $s = j\omega_k$ and equating to the response measured for the angular frequency ω_k, we get

$$\sum_{i=1}^{r} H_i(j\omega_k) = p_k + jq_k \qquad (k = 1, \cdots, m/2) \qquad (28)$$

Each rational polynomial H_ρ can be identified over the corresponding response section by using Gauss-Seidel type iterations given by

$$H_\rho^l(j\omega_\kappa) = p_\kappa + jq_\kappa - \sum_{\substack{i=1 \\ i\neq\rho}}^{r} H_i^{l-1}(j\omega_\kappa) \qquad (\rho = 1,...,r) \qquad (29)$$

Here l denotes the iteration step and ω_κ the observation frequencies between the limits for the corresponding section. All parameters can initially be assumed zero and at each iteration step, the H_ρ terms are identified in terms of the corresponding observation data and the remaining terms with parameters obtained at the last step. Numerical examples have shown that the iterations converge to minimize the residual, provided that an appropriate order has been chosen by a priori knowledge or trials.

Iterative Improvement

The results obtained with the least squares approach, either by direct solution or partitioning, may be further improved by Gauss-Newton refinement. The need for this refinement arises for two reasons. First, the results may still have large errors due to the finite digit accuracy of the calculations and inevitable measurement errors combined with the ill-conditioning of the problem. Second, linearization of (2) has resulted in the badly distorted linear system (6), due to the undesired weighting produced by the multiplication with the denominator of the original problem. Moreover, matrix A of (6) contains in half of its columns the measured data p_k, q_k, so that the problem actually resembles the more comprehensive *Total Least Squares* (TLS) problem analyzed in [11]. Since the solution process does not contain any physical constraint, right-half plane poles may be obtained for systems which are known to be stable. Although such a solution is mathematically acceptable, it is not meaningful and obviously not applicable for engineering purposes. Let us assume that the coefficients of the transfer function have been calculated by the solution of the least squares problem as described in the previous sections. Since the polynomial coefficients are known, the poles can be calculated as

$$\lambda_i = \sigma_i + j\omega_i \qquad \lambda_i^* = \sigma_i - j\omega_i \qquad (30)$$

Here, all the poles are assumed complex conjugate for generality. A simple real pole can be considered as a special case where the corresponding imaginary part is zero, and the conjugate is ignored. The transfer function can be written in terms of partial fractions as

$$H(s) = \sum_{i=1}^{v/2} \left[\frac{\alpha_i + j\beta_i}{s - \sigma_i - j\omega_i} + \frac{\alpha_i - j\beta_i}{s - \sigma_i + j\omega_i} \right] = \sum_{i=1}^{v/2} H_i(s) \qquad (31)$$

By substituting $s = j\omega_k$ into (31) and using the observation data corresponding to ω_k we can write the equation

$$f_k = \sum_{i=1}^{v/2} \left[\frac{\alpha_i + j\beta_i}{-\sigma_i + j(\omega_k - \omega_i)} + \frac{\alpha_i - j\beta_i}{-\sigma_i + j(\omega_k + \omega_i)} \right] - p_k - jq_k \approx 0 \qquad (32)$$

For $m/2$ different values of ω_k, (32) represents a set of overdetermined

nonlinear equations, which may be solved by Newton's method (Gauss-Newton, because of the least squares nature of the problem) [12] :

$$\mathbf{J}_{old}(\xi - \xi_{old}) = -\mathbf{f}(\xi_{old}) \qquad (33)$$

where,

$$\xi^T = [\cdots \alpha_i, \beta_i, \sigma_i, \omega_i \cdots] \qquad (i=1, \cdots, v/2)$$
$$\mathbf{f} = [f_1, f_2, \ldots, f_{m/2}]^T = \mathbf{f}' + j\mathbf{f}'' \qquad (34)$$

$\mathbf{J}_{old} = \mathbf{J}' + j\mathbf{J}''$ is the complex jacobian matrix of $\mathbf{f}$ evaluated in terms of the parameters computed at the last iteration. Defining a new set of variables $\Delta\xi = \xi - \xi_{old}$ and separating the real and imaginary parts, (33) can be written in the general form of a linear least squares problem with real coefficients:

$$\Psi \; \Delta\xi = \phi$$
$$\Psi = \begin{bmatrix} \mathbf{J}' \\ \mathbf{J}'' \end{bmatrix} \qquad \phi = \begin{bmatrix} \mathbf{f}' \\ \mathbf{f}'' \end{bmatrix} \qquad (34)$$

Equation (34) can be solved by SVD with the *inequality constraints* $\sigma_i \le 0$, simply enforced on the SVD solution of (34), which eliminates the possibility of obtaining right plane poles. Besides, (34) is not weighted as (6) or (17), so that the undue emphasis on the observations corresponding to the frequencies near the zeros of the transfer function will be overcome by the iterative improvement.

General Algorithm for the Estimation Process

The complete method for estimation of the transfer function from frequency response observation data can be summarized by the following algorithm :

1 – Read frequency response data and the tolerance level for the residual

2 – Partition the frequency scale by considering the form of the observation data (If the partitioning is not possible or necessary, take the whole range as a single section; this is a special case where $r=1$)

3 – Estimate the order of the partial terms

4 – Form the set of equations as described by expression (6)

5 – Perform column scaling, check the condition number

6 – Identify the parameters of each partial term by the Gauss-Seidel iterations (In the special case of a single section, a single step will give the solution directly)

7 – If the residual is below the tolerance level and all the poles are in the left half-plane, stop the process

8 – Perform iterative improvement until the residual becomes smaller than the given tolerance.

Examples

The presented method of estimation has been applied to a number of examples, to test its effectiveness and accuracy for different practical cases. For simplicity, the frequency response data were generated by using mathematical models with known poles, λ_i, λ_i^*, and pole strengths, $\alpha_i \pm j\beta_i$. In the case of $v/2$ complex conjugate pairs of poles, the frequency response can be calculated from

$$H(j\omega) = \sum_{i=1}^{v/2} \left[\frac{\alpha_i + j\beta_i}{j\omega - \lambda_i} + \frac{\alpha_i - j\beta_i}{j\omega - \lambda_i^*} \right] \qquad (35)$$

In this section some selected examples will be presented for the purpose of illustration.

Description of the model systems

The model systems referred to as M1 – M8 in Table 1, of different levels of complexity, were used for numerical examples. M1, M2, and M3 are second order systems with the same $\lambda_{imag}/\lambda_{real}$ ratio, to ensure similar responses with respect to the normalized frequency $\overline{\omega} = \omega/\lambda_{imag}$. The pole strengths are chosen as $\alpha = \lambda_{real}$ and $\beta = 0$ to have peak values close to the unity. This particular selection results in approximately equal observation values p_k, q_k in the normalized observation range. Each set of response data consists of the real and imaginary parts of the transfer function, observed at 100 discrete ω values, equally spaced along the logarithmic frequency scale. M4 and M5 are fourth order systems formed as combination of M1, M2 and M2, M3, respectively. M6 is a sixth order system created by combining M1, M2, M3, with a larger observation interval. M7 and M8 are derived from M6 by moving the pole with the largest imaginary part parallel to the imaginary and real axes.

Condition numbers

The condition numbers calculated by (15) for the described models are presented in Table 1. They increase with the model order and the maximum observation frequency. Numerous examples have shown that the condition number increases as ω_{max}^v, where ω_{max} is the maximum observation frequency and v is the model order. The improvements achieved by the row and column scaling are also shown in the table. It can be concluded that the column scaling is significantly more effective than the row scaling. However, there are cases when the condition number may still be in the order of 10^7 or even larger. It is therefore suggested to use double precision arithmetic to avoid excessive computational errors in such cases.

The estimation procedure

The model M8 described in Table 1 will be used as an example for the application of the procedures discussed in the previous section. The magnitude of the frequency response obtained by using (35) is plotted in Figure 1. A realistic prediction of the model order is important for the accuracy of the results. The form of the response curve reveals the existence of 3 groups of poles. This curve can be partitioned and Gauss-Seidel iterations can be used for identification of three rational polynomials. However, in this particular example the data is generated by computation, rather than measurements, and the error level is very low. In such cases, the error amplification due to the poor condition number remains below reasonable tolerance limits, so that estimation without partitioning may be successful. As the form of the response curve shows that the system has at least 3 complex conjugate poles, the condition numbers (with column scaling!) and the output errors were calculated and listed in Table 2 for several, different model orders.
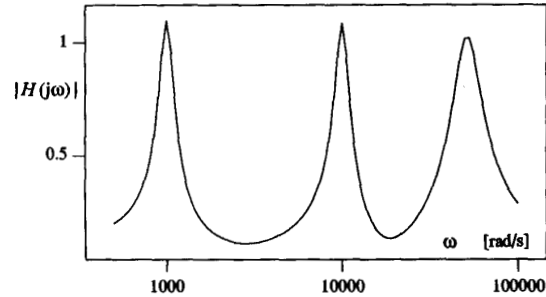

Figure 1 Frequency response of M8

Table 1 Example models and condition numbers

Model identifier	Model order v	Pole locations		Observation range		Condition number		
		Real part	Imag. part	ω_{min}	ω_{max}	Without scaling	Row scaling	Column scaling
M1	2	−100	±1000	500	1500	0.62×10^6	0.51×10^5	2.3
M2	2	−1000	±10000	5000	15000	0.62×10^9	0.51×10^7	2.3
M3	2	−2000	±20000	10000	30000	0.25×10^9	0.21×10^8	2.3
M4	4	−100 −1000	±1000 ±10000	500	15000	0.58×10^{16}	0.12×10^{12}	376.8
M5	4	−1000 −2000	±10000 ±20000	5000	30000	0.18×10^{18}	0.22×10^{16}	25.7
M6	6	−100 −1000 −2000	±1000 ±10000 ±20000	500	30000	0.20×10^{27}	0.36×10^{22}	0.15×10^5
M7	6	−100 −1000 −2000	±1000 ±10000 ±50000	500	100000	0.61×10^{28}	0.15×10^{22}	0.67×10^6
M8	6	−100 −1000 −10000	±1000 ±10000 ±50000	500	100000	0.23×10^{30}	0.48×10^{22}	0.25×10^7

Table 2 Condition number and error for different model orders

Order	Cond. No.	Error (%)
4	0.28×10^5	17.44
5	0.24×10^5	16.38
6	0.25×10^7	0.013
7	0.20×10^9	0.008
8	0.25×10^9	0.003

Table 3 Identified parameters for M8

s^k	Numerator	Denominator
s^0	0.1168	1.0000
s^1	0.2469×10^{-3}	0.2255×10^{-3}
s^2	0.1082×10^{-6}	0.1005×10^{-5}
s^3	0.3124×10^{-10}	0.2936×10^{-10}
s^4	0.1096×10^{-14}	0.1035×10^{-13}
s^5	0.8368×10^{-19}	0.8368×10^{-19}
s^6		0.3769×10^{-23}

In the case of an under-estimated model order (4 or 5), the output error is quite high. For model order 6, which is equal to the actual order, the error drops to a very low level. In the case of over-estimated model orders, the output error becomes even lower. Examination of the results obtained for over-estimated model orders shows that some of the parameters are negligible, so that the identified polynomials reflect in fact the actual order. The results of the non-partitioned estimation (the whole frequency range is taken as a single section) are presented in Table 3. Note that the coefficients appear to be small only if viewed in isolation but are significant if combined with the respective powers of ω (for instance $\omega = 10^4$).

The actual pole locations for M8 are given in Table 1. The pole strengths are all real and equal to the real part of the corresponding pole. The pole locations and strengths calculated from the estimated rational polynomial (specified by Table 3) are given in Table 4.

Table 4 Pole locations and strengths calculated from the estimated model

Pole location	Pole strength
$-0.99999988 \times 10^4 \pm j0.50000000 \times 10^5$	$0.99999986 \times 10^4 \pm j0.10313583 \times 10^{-2}$
$-0.99998537 \times 10^3 \pm j0.10000001 \times 10^5$	$0.99995927 \times 10^3 \pm j0.17150874 \times 10^{-1}$
$-0.10003031 \times 10^3 \pm j0.10001574 \times 10^4$	$0.10004375 \times 10^3 \mp j0.13023601$

The results presented above show that both the frequency response and the poles calculated from the estimated transfer function are in a good match with the actual system. However, it should be noted that the input data, generated by computation were accurate to eight digits. Due to this high accuracy of the input data, the error in the output, obtained by double precision computation, is small despite the high condition number (in the order of 10^7). In the cases where measurement errors exist in the input, partitioning is needed to increase the accuracy of the estimation, since the condition number corresponding to each partial term will then be much smaller. Examples for such cases are given in the section related to the representation of transformers.

State Equation Representation of Estimated Transfer Function

Once the transfer function is identified as explained above, the state equations can easily be written in terms of the obtained parameters. Suppose that the transfer function has been identified as a sum of r partial fractional terms,

$$H(s) = \sum_{i=1}^{r} \left[\frac{a_{0i} + a_{1i}s + \cdots + a_{\mu i}s^{\mu i}}{1 + b_{1i}s + \cdots + b_{vi}s^{vi}} \right] \quad (36)$$

It is assumed that the systems considered are strictly proper, and $\mu = v - 1$. The transfer function (36) can be directly represented by state equations, using the companion form for each rational polynomial [13] :

$$\dot{z} = Fz + gu$$
$$y = c^T z \quad (37)$$

where,

$$F = \text{diag}\{F_i\} \quad g = \text{col}\{g_i\} \quad c = \text{col}\{c_i\}$$

$$F_i = \begin{bmatrix} 0 & 1 & 0 & . & 0 \\ 0 & 0 & 1 & . & 0 \\ 0 & 0 & 0 & . & . \\ . & . & . & & 1 \\ -\frac{1}{b_{vi}} & -\frac{b_{1i}}{b_{vi}} & -\frac{b_{2i}}{b_{vi}} & . & -\frac{b_{\mu i}}{b_{vi}} \end{bmatrix} \quad g_i = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ \frac{1}{b_{vi}} \end{bmatrix} \quad c_i = \begin{bmatrix} a_{0i} \\ a_{1i} \\ . \\ . \\ a_{\mu i} \end{bmatrix} \quad (38)$$

The state equations (37) contain only the parameters resulting from the transfer function estimation, and no additional computation is needed to obtain the state space representation for the given system. This form of the state equations is general and applicable to the different procedures presented in this paper. The non-partitioned (single section) estimation corresponds to the special case where $r = 1$. In the case when iterative improvement is applied, the transfer function is obtained as a sum of quadratic terms, yielding 2×2 blocks in F.

APPLICATION TO REPRESENTATION OF TRANSFORMERS

The need and importance of wide frequency range representation of power transformers have been discussed in a number of papers [14]-[17]. In practice, this type of representation may basically be of interest for the study of electromagnetic transients in power transmission networks. An accurate representation of transformers is important for correct prediction of possible system overvoltages that may occur in transient conditions such as switching operations, propagation of lightning surges or some types of faults. The estimation method presented in this paper can be applied for obtaining the transfer function or state equation representation of a transformer winding described only by external frequency response measurements. At this stage, the single-input, single-output case is considered. Extension of the method to multiple input and output applications will be covered in future work.

The frequency response of transformers is usually determined by wide frequency range terminal impedance or admittance measurements. The equipment and procedures for such measurements are described in [17]. The available test systems with data-acquisition facility can measure, process, and store a large number of observation data from a very low frequency range, below the industrial frequency, to the high MHz level. These observations form a convenient data base for the estimation procedure presented in this paper.

The general features of the high frequency behavior of transformers are presented and discussed in [18]. The test results have shown that the frequency response depends significantly on the terminal (loading) conditions. In general, the frequency response curve of a transformer may have several peaks, resulting from resonance phenomena. At the lower frequency range the impedance value is very small compared to the peak values, due to the small winding resistance. At the high frequency range, the capacitances shunting the winding sections are more dominant, so that the impedance approaches to zero as the frequency reaches very high values, that may practically be considered as infinity. A transformer can therefore be considered as a strictly proper system, and represented by an impedance transfer function as in (1).

The application of the suggested estimation procedure to the representation of transformers is demonstrated by using sets of frequency response data measured for different type of transformers. A brief description of the physical properties of these transformers is given below.

Transformer-1

This example is a laboratory size single phase model transformer, of approximately 0.5 kVA. The windings are dry insulated and, due to the small dimensions, the inductances of winding sections are more predominant than the shunt and ground capacitances. The magnitude of the terminal impedance measured with the secondary winding open is shown in Figure 2 by a dashed line.

Transformer-2

Transformer-2 is a 3 phase, Δ/Y connected, 75 kVA, 4160/208 V oil filled distribution transformer. The impedance characteristic is measured from the HV side, the terminals H1 and H2 connected together, and H3 grounded. The measurement is made between the common terminal H1-H2 and H3 (ground) while the low voltage side is shorted. The magnitude of the impedance is shown by a dashed line in Figure 4.

Applied Estimation Methods

The frequency response data obtained for the transformers described above are used for estimation of the transfer functions. In the case of *Transformer-1*, the non-partitioned (single section) estimation is not successful. The response from a 6th order estimated transfer function is shown in Figure 2 by a solid line. The error amplification is so high that the resonant behavior of the actual winding is not reflected by the estimated model.
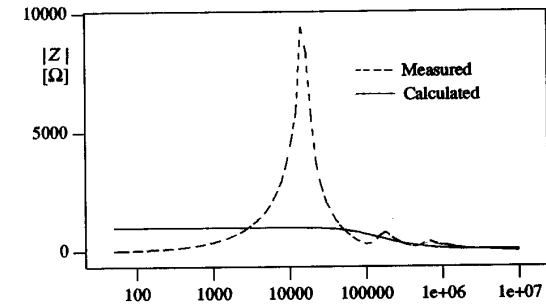

Figure 2 Frequency response of Transformer-1 and the result of the non-partitioned estimation.
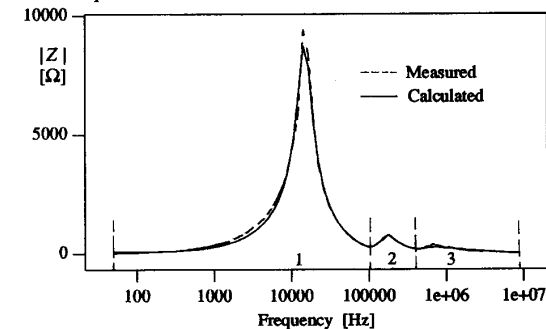

Figure 3 Result of the partitioned estimation for Transformer-1

To increase the accuracy, the frequency scale was partitioned as shown in Figure 3. The partitioned estimation followed by iterative improvement yielded an RMS error of 2.65% in the magnitude. The frequency response calculated from the estimated transfer function is drawn as a solid line in Figure 3.

The numerical results of the estimation for Transformer-1 are given in Table 5.

Table 5 Identified parameters for Transformer-1

Sequence	s^k	Numerator	Denominator
1	s^0	0.5588×10^2	1.0000
	s^1	0.4300×10^{-1}	0.9678×10^{-5}
	s^2	0.1578×10^{-6}	0.1397×10^{-9}
	s^3	0.1483×10^{-12}	0.5957×10^{-15}
	s^4	0.2013×10^{-18}	0.3554×10^{-21}
	s^5		0.6974×10^{-27}
2	s^0	-0.1981×10^3	1.0000
	s^1	0.1671×10^{-3}	0.5867×10^{-6}
	s^2	-0.8032×10^{-10}	0.1102×10^{-11}
	s^3	0.2246×10^{-16}	0.2707×10^{-18}
	s^4		0.8963×10^{-25}
3	s^0	0.2219×10^3	1.0000
	s^1	0.1336×10^{-5}	0.1528×10^{-6}
	s^2	0.1049×10^{-12}	0.3043×10^{-15}
	s^3		0.1260×10^{-22}

The frequency response curve obtained for Transformer-2 is shown in Figure 4 by a dashed line. With this data, the estimation has been successful without partitioning. However, the unconstrained least squares estimation yielded right half-plane poles, as shown in Table 6. These unstable poles were then eliminated by iterative improvement. The new pole locations and strengths are given in Table 7. The RMS error on the magnitude, drawn with a solid line in Figure 4, is 1.08%, which is acceptable for most applications.
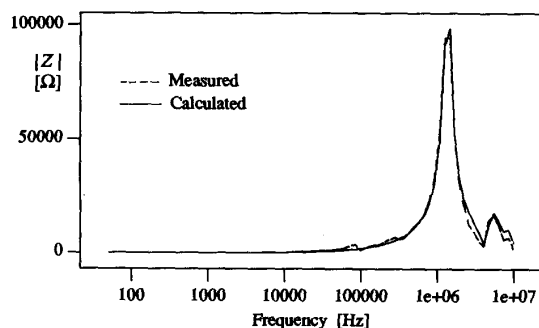

Figure 4 Result of the estimation with iterative improvement for Transformer-2.

Table 6 Pole locations and strengths for Transformer-2 without iterative improvement.

Pole location	Pole strength
$-0.3839 \times 10^7 \pm j0.3083 \times 10^8$	$0.4801 \times 10^{11} \pm j0.3867 \times 10^{11}$
$-0.9423 \times 10^6 \pm j0.8498 \times 10^7$	$0.1057 \times 10^{12} \pm j0.1804 \times 10^{11}$
$0.9737 \times 10^6 \pm j0.3998 \times 10^8$	$-0.5284 \times 10^{10} \pm j0.3799 \times 10^{10}$
$0.1680 \times 10^7 \pm j0.5637 \times 10^8$	$-0.9080 \times 10^{10} \pm j0.2436 \times 10^{10}$
0.8774×10^8	-0.7319×10^{12}

Table 7 Pole locations and strengths for Transformer-2, after iterative improvement.

Pole location	Pole strength
$-0.3588 \times 10^7 \pm j0.3054 \times 10^8$	$0.4036 \times 10^{11} \pm j0.3923 \times 10^{11}$
$-0.9424 \times 10^6 \pm j0.8445 \times 10^7$	$0.1048 \times 10^{12} \pm j0.2342 \times 10^{11}$
$-0.1453 \times 10^7 \pm j0.3992 \times 10^8$	$-0.7765 \times 10^{10} \mp j0.5574 \times 10^{10}$
$-0.2409 \times 10^7 \pm j0.5668 \times 10^8$	$-0.1289 \times 10^{11} \mp j0.2803 \times 10^{10}$
-0.2446×10^8	0.1151×10^{12}

CONCLUSIONS

A methodology for the estimation of a transfer function from frequency response measurements is presented and discussed. The transfer function parameters can be identified by the solution of an overdetermined set of equations in the least-squares sense. The formulation, by its nature, leads to a poorly conditioned problem. The literature abounds in examples of problems that often could not be solved due to their poor numerical condition. The methodologies described in this paper will in general overcome the difficulties that have previously been reported.

The following are the main results of the paper.

- The condition number increases mainly with the order of the transfer function, and the maximum observation frequency. In many cases, the condition number can be excessively large and the solution of the equations becomes impossible even with the double precision computation. Three approaches, termed *shifting, row,* and *column scaling* have been introduced to improve the condition of the problem. Though all of these methods reduced the condition number, the *column scaling* provided the only truly significant improvement.

- *Singular Value Decomposition* has been used for the solution of the overdetermined set of linear equations, pre-processed by column scaling. When the error level in the input data is relatively small, for example, in the case of data generated by computation, it is possible to perform the estimation with very low output error, for a large number of poles. However, in practical cases where measurement errors exist in the input data, or the response is obtained from a distributed parameter system, as in the case of transformers, the direct estimation may not be successful with the required degree of accuracy. This difficulty can be overcome by *partitioning* the frequency scale and applying Gauss-Seidel iterations.

- In some cases, the arbitrariness of the least squares solution may lead to right half-plane poles for systems which are known to be stable. An *iterative improvement* method with constraints on the real part of the poles can be applied both to avoid the possibility of having unstable poles and to increase the accuracy of the estimation.

The application of the proposed procedures has been illustrated by different examples. In the case of computer generated input data, non-partitioned estimation can often be expected to be successful with an output error of less than 0.01%. Two examples are given for representation of transformers. In one of these examples, an acceptable estimate could not be obtained without partitioning. After partitioning, the estimation has been successful, with an output error below 3%. The last example illustrates the necessity of the iterative improvement. The direct solution of the least squares problem yielded unstable poles which have subsequently been eliminated by the iterative improvement.

While the chosen examples for applications were from the group of distribution transformers, because of the availability of measurement data, the estimation methodologies described in the paper could readily be applied to large power transformers. In fact their applicability extends beyond the general field of transformers used here for the purpose of demonstration.

All procedures described apply to scalar transfer functions. The extension of these methods to multi-input and output problems will be presented in a sequel to this paper.

REFERENCES

[1] E.C. Levy, "Complex Curve Fitting", IRE Trans. on Automatic Control, Vol. AC-4, pp. 37-44, May 1959.

[2] C.K. Sanathanan and J. Koerner, "Transfer Function Synthesis as a Ratio of Two Complex Polynomials ", IEEE Trans. on Automatic Control, Vol. 8, pp. 56-58, Jan. 1963.

[3] P.A. Payne, "An Improved Technique for Transfer Function Synthesis from Frequency Response Data", IEEE Trans. on Automatic Control, Vol. 14, pp. 480-483, Aug. 1970.

[4] P.J. Lawrence and G.J. Rogers, "Sequential Transfer-Function Synthesis from Measured Data", Proc. IEE, Vol. 126, No. 1, pp. 104-106, Jan. 1979.

[5] H. Stahl, "Transfer Function Synthesis Using Frequency Response Data", Int. Journal of Control, Vol. 39, No. 3 pp. 541-550, 1984.

[6] A.H. Whitfield, "Transfer Function Synthesis Using Frequency Response Data", Int. Journal of Control, Vol. 43, No. 5, pp. 1413-1426, 1986.

[7] C.L. Lawson and R.J. Hanson, *Solving Least Squares Problems*, Prentice-Hall, Inc. Englewood Cliffs, N.J., 1974.

[8] G.E. Forsythe, M.A. Malcolm, and C.B. Moler, *Computer Methods for Mathematical Computations*, pp. 41-48, 96, Prentice-Hall, Inc. Englewood Cliffs, N.J., 1977.

[9] A. Van der Sluis, "Condition Numbers and Equilibration of Matrices" Numer. Math., Vol. 14, pp. 14-23, 1969.

[10] A. Van der Sluis, "Stability of Solutions of Linear Algebraic Systems", Numer. Math., Vol. 14, pp. 246-251,1970.

[11] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem - Computational Aspects and Analysis*, SIAM, 1991.

[12] R. L. Johnston, *Numerical Methods - A Software Approach*, John Wiley and Sons Inc., New York, 1982.

[13] T. Kailath, *Linear Systems*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1980.

[14] W. J. McNutt, T. J. Blalock, and R. A. Hinton, "Response of Transformer Windings to System Transient Voltages", IEEE Trans. on Power Apparatus and Systems, Vol. PAS-93, pp. 457-467, March-April 1974.

[15] R.J. Musil, G. Preininger, E. Schopper, and S. Wenger, "The Resonance Effect of Oscillating System Overvoltages on Transformer Windings", IEEE Trans. on Power Apparatus and Systems, Vol. PAS-101, pp. 3703-3711, Oct. 1982.

[16] R. C. Degeneff, M. R. Gutierrez, and P. J. McKenny, "A Method for Constructing Reduced Order Transformer Models for System Studies from Detailed Lumped Parameter Models", Proc. IEEE Transmission and Distribution Conference, pp. 532-538, 1991.

[17] R. Caldecott, Y. Liu, S. A. Sebo, D. G. Kasten, and S. E. Wright, "Measurement of the Frequency Dependent Impedance of Major Station Equipment", IEEE Trans. on Power Delivery, Vol. 5, No. 1, pp. 474-480, January 1990.

[18] A. O. Soysal, "A Method for Wide Frequency Range Modeling of Power Transformers and Rotating Machines", Proc. IEEE Transmission and Distribution Conference, pp. 560-566, 1991.

A. Oguz Soysal (M'90) was born in Turkey, in 1951. He received the B.Sc, M.Sc. and Ph.D degrees from Istanbul Technical University. In 1983 he joined ESAS Electrical Industry-power transformer factory as R/D engineer. In 1986 he was appointed Assistant Professor at Karadeniz Technical University, Turkey. At present, Dr. Soysal is with the same institution as Associate Professor and is spending his sabbatical leave at the University of Toronto. His research areas include analysis and control of power systems, power transmission and distribution, study of electromagnetic transients and system identification.

Adam Semlyen (F'88) was born and educated in Rumania where he obtained a Dipl. Ing. degree and his Ph.D. He started his career with an electric power utility and held an academic position at the Polytechnic Institute of Timisoara, Rumania. In 1969 he joined the University of Toronto where he is a professor in the Department of Electrical Engineering. His research interests include the steady state and dynamic analysis of power systems, electromagnetic transients, and power system optimization.

Discussion

H. Tsai and **A. Keyhani** (The Ohio State University, Electrical Engr., Columbus, OH 43210): We would like to commend the authors for a well-written paper and for their efforts to develop practical methodology for high frequency modeling of electrical power apparatus, particularly the transformers.

As shown by the authors, the formulation of the transfer function estimation in the sense of linear least squares can significantly reduce the complexity of the estimation problem. Due to the ill conditioning of the problem, scaling of the estimated parameters and partitioning of the transfer function models are necessary for the estimation. Using these two remedies, a nonlinear least squares estimation (NLSE) method can also be used to establish the transfer function models for high frequency modeling of transformer windings [1]. When using the NLSE technique, the transfer function model is represented by real and complex poles and zeros. Therefore, the anti-resonant and resonant frequencies of the frequency responses can be directly represented in the structure of the transfer function model. For nonlinear estimation, iterative procedure is required where the unknown parameters need to be initialized. This can be done by directly checking the frequency response measurement for the critical points.

Another important aspect of any system identification procedure from the experimental data is the noise effect on the measured responses. For noisy data, linear least squares estimation may not produce unbiased estimation results with minimal error covariance. Therefore, it is our opinion that certain measure should be considered in the estimation process for taking care of the noise in the data [1].

As shown in Eq. (7) of the paper, the matrix $A^T A$ is not a sparse matrix. As more resonant conditions exist in the frequency spectrum, the order of the transfer function as well as the size of the matrix increase. From Table 1 of the paper, the highest order model used in the study was six. Furthermore, from Fig. 1 of the paper, the assumed transfer function model seems to have no complex zeros, or the terminal resonant conditions. Was it done in order to reduce the complexity of the problem at hand? The technique introduced in this paper is computationally simpler compared to other nonlinear methods. Then is it possible to use this technique to identify a much higher order transfer function model than the sixth order model considered in this paper?

From Fig. 2 and Fig. 3, the phase comparison of the measured and the simulated responses is not provided. We believe that the phase information is important because it is used in the estimation process and can be easily provided in the comparison study.

The authors have provided the power industry with a valuable and practical technique for estimating the parameters of transformer winding high frequency transfer functions. We would appreciate the author's comments concerning the questions and issues raised in this discussion.

Reference

[1] A. Keyhani, H. Tsai, and A. Abur, "Maximum likelihood estimation of high frequency machine and transformer winding parameters," IEEE Trans. on Power Delivery, Vol. 5, No. 1, Jan. 1990, pp. 212-219.

A. S. Morched and **L. Martí** (Ontario Hydro, Toronto, Canada): The authors should be congratulated for presenting a thorough and detailed analysis of problems related to transfer function estimation from frequency domain data. The authors have also presented a procedure to improve the solution of a fairly ill-conditioned problem.

In earlier work, we also encountered similar difficulties in the fitting of rational functions to measured or computed transfer functions. The implementation of techniques similar those described in the paper resulted in major improvements in the accuracy of the calculations. However, these techniques, by themselves, were not enough to resolve system singularities in several instances:

- High-order systems.
- Systems with closely packed and/or highly damped poles.
- Transfer functions with high noise content and/or background poles.

One way to alleviate these problems was to use Chebyshev polynomials, rather than a power series, for the fitting of rational functions [A]. Power series have very large dynamic range when ω is not normalized (in the order of ω^n), and their shapes (as a function of n) are very similar when ω is normalized, as shown in Figure 1. This behavior contributes to ill-conditioning in least-square fitting calculations.

In contrast, Chebyshev polynomials always have a small dynamic range (between 1 and -1) and each polynomial has a distinctive shape when ω is normalized, as shown in Figure 2. This makes Chebyshev polynomials particularly well suited to least squares fitting.

After solving for the coefficients of the numerator and denominator of the transfer function expressed with Chebyshev polynomials, the equivalent coefficients of the power series can be formed. Have the authors considered using Chebyshev polynomials (or other orthogonal polynomials) to improve the accuracy of the fitting process?

One important feature of the least squares process is the use frequency dependent weighting functions. Have the authors considered using weighting functions linked to the location of the poles and zeroes of the transfer function?

Lastly, it is not immediately obvious from the paper how the authors determine the order of the system. Is this done by inspection? Have the authors developed an automatic/systematic procedure to assess the order of the approximation?

The comments of the authors on these points will be much appreciated.

Reference

[A] James L. Adcock, "Curve Fitter for Pole-Zero Analysis," *Hewlett Packard Journal*, Jan. 1987. pp. 33-36.
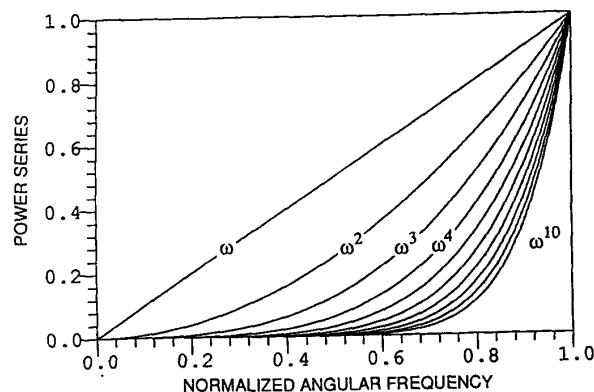
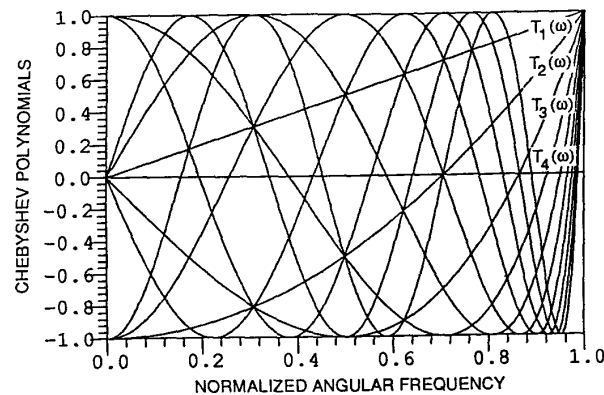**Fig 1:** *Power series.*

Fig 2: *Chebyshev polynomials.*

R. Malewski (Westmount, Quebec, Canada): The authors should be congratulated for their study on circuit synthesis from experimental data. Among several potential users the transmission system planners are interested in a high frequency model of transformer. Such model is required to predict, for instance, the transfer of transient overvoltages from one voltage level to another. From the utility point of view the simulation of large HV power transformers is of main interest. An application of the presented method to say 800 kV class, 500 MVA unit would attract attention of the planning department of large utilities.

Obviously, measurements of the transformer transfer function can not be taken on such a large unit at the University facilities, but the actual measured characteristics can be obtained from an industrial laboratory performing the acceptance test of new transformers. At present, many laboratories use a digital recorder for monitoring the impulse test. The obtained records are processed in order to enhance the efficiency of fault detection. The processing often includes calculation of the frequency spectrum of the output and input impulses, and finding the transformer transfer function as quotient of these two spectra. An analysis of the transfer function required for the dielectric fault detection, is not pertinent to the study presented by the Authors. However, a large pool of experimental data accumulated on large HV power transformers can be used for the transformer circuit synthesis and development of the computer model.

The routine test procedure calls for measurement of the applied test impulse, and of the output current at the winding neutral terminal. These records yield the winding transadmittance, whereas the transfer function between two winding may be more important for the circuit synthesis. However, at a reduced level, additional records can be taken during the impulse test, if requested by the utility purchasing the transformer. Such additional measurements can be included in the test program, on demand of the utility system planning department. An incremental cost of the additional measurement is negligible, since the impulse generator and recording system are anyhow prepared for the acceptance test.

W.-G. Huang (Department of Electrical Engineering, Tsing-hua, University, Beijing, P. R. China): The discusser would like to compliment the authors for presenting a methodology for the estimation of the transfer function parameters from terminal frequency response measurements, and for overcoming the difficulties encountered in the parameter identification due to the poor numerical condition. Column scaling has been efficiently used to improve the condition of the problem. In most of the cases when the model order is high, though pre-processed by column scaling, the error amplification is still too large. This difficulty has been overcome by partitioning the frequency scale and applying Gauss-Seidel iterations. In some cases unstable right-half plane poles have been eliminated by solving the least squares equations with inequality constrains and the accuracy of estimation has been increased by an iterative improvement method. These novel procedures proposed in the paper are superior to existing ones and are very instructive.

The ill condition problem in the paper reminds the discusser of the pathological phenomenon of the model transformation matrix $[T]$ of the cable/GIS bus propagation transfer function matrix $[P]$ at high frequency. When three phase cables are unsymmetrically buried under the ground (6 by 6 parameter matrices) or the cores of the three phase GIS bus are unsymmetrically arranged inside the sheath (4 by 4 parameter matrices) and the frequency is high, three eigenvalues of the propagation transfer function matrix $[P]$ are very close to each other. Small pertubation on the cable/GIS parameters will cause significant change of the eigenvectors related to these three eigenvalues, and the model transformation matrix $[T]$ will be changed significantly. The problem has been solved as follows [1].

Matrix $[D]$ is introduced to transfer the equation $[P][T] = [T][\Lambda]$, where $[\Lambda]$ is a diagonal matrix, the elements of which are eigenvalues of $[P]$.

The equation can be written as $[D]^{-1}[P][D][D]^{-1}[T] = [D]^{-1}[T][\Lambda]$ with $[P'] = [D]^{-1}[P][D]$, and $[T'] =[D]^{-1}[T]$, then $[P'][T'] = [T'][\Lambda]$.

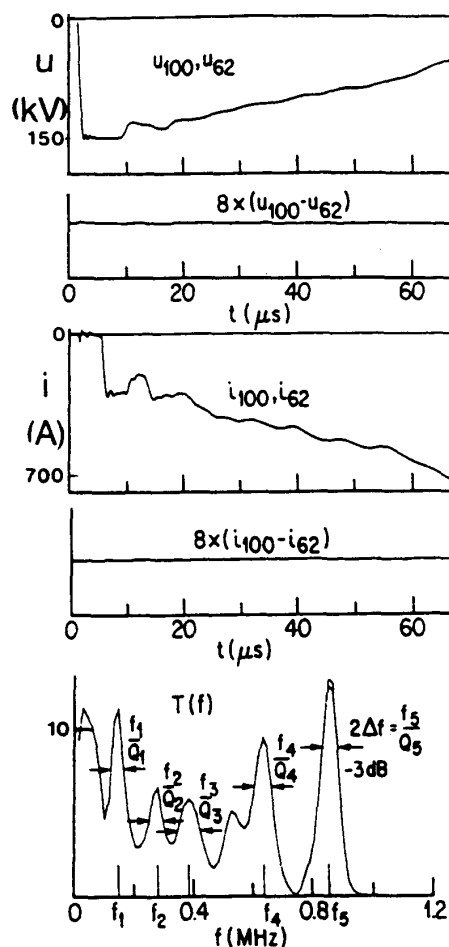$$[D] = \begin{pmatrix} [I] & [I] \\ [I] & [0] \end{pmatrix} \qquad \text{for cable,}$$

Fig. 1. A typical record taken during the impulse test of a large HV power transformer. Records of the test impulse applied at the full (100%) and reduced (62%) basic insulation level are superimposed in upper graph. Their difference is shown below, magnified by factor 8.

The neutral terminal current records corresponding to the full and reduced level applied impulse (100% and 62% respectively) are shown in middle graph, together with their eightfold magnified difference.

The winding transadmittance (frequency spectrum of the neutral current divided by the applied voltage spectrum) is plotted in lower graph. A number of resonant frequencies can be seen on this "transfer function." This is a typical behavior of a layer type winding.

$$[D] = \begin{pmatrix} [I] & \begin{matrix} & 1 \\ & 1 \\ & 1 \end{matrix} \\ 0\ 0\ 0 & 1 \end{pmatrix} \qquad \text{for GIS bus,}$$

where $[I]$ is an identity matrix.

After such transformation, $[T']$, thus $[T]$, can be solved reasonably.

$$[T] = \begin{pmatrix} [T2] & [T1] \\ [T2] & [0] \end{pmatrix} \qquad \text{for cable,}$$

$$[T] = \begin{pmatrix} 1 & \\ 1 & [T2] \\ 1 & \\ 1 & 0\ 0\ 0 \end{pmatrix} \qquad \text{for GIS bus.}$$

If the arrangement of cables or the cores of GIS bus is symmetrical, Clarke or Karenbauer transformation matrix can be used as submatrix $[T2]$.

Any comment concerned with curing the pathological problem related to the model transformation matrix $[T]$ of the unsymmetrically arranged cable/GIS bus from the authors' experience is very appreciated.

Wide frequency range modeling of major station equipment, such as transformers transmission lines, reactors etc., is sometimes needed for the radio frequency electromagnetic noise and the overvoltage calculations.

A transformer sometimes is represented by a series of four terminal networks built up by self inductances (L), mutual inductances (M), transversal capacitances (C) and longitudinal capacitances (K) of the winding (LMCK model for short), voltages in the winding and voltage gradients between sections or turns can easily be calculated. If the lightning overvoltage on the terminal of a transformer is of interest, usually a transformer is represented by its input equivalent capacitance. The wide frequency range model is a more accurate "black box" model, the parameters of the transfer function of which can be estimated by external frequency response measurement. According to our experience, when the operating voltage is neglected (zero initial condition), no matter whether the transformer is represented by a LMCK equivalent or by an equivalent capacitance the lightning overvoltages on the transformer terminal and on other apparatus in the substation protected by current-limiting gap arresters are approximately unchanged. With either of the equivalent the difference of the overvoltages on the transformer terminal is less than 3%. If the operating voltage is taken into account (non-zero initial condition) the difference will be ranged over 13% to 18% [2]. The computer program of the LMCK equivalent of a transformer winding has been completed by eliminating mutual inductances (M) and by using inductance as basic element of the discretized equivalent circuit, and the program interfaces with EMTP, surges in transformer winding and on other apparatus under operating condition can be studied simultaneously [2]. The discusser is curious to know whether the wide frequency range model is easy to be implemented in or to be interfaced with EMTP, and to see what would be the results compared with other two models in the overvoltage calculation under operating condition in substations.

References

[1] Wang Hong-Li, A study of Cable/GIS Transient Parameter Calculation. doctorate dissertation, Tsing-ha University, Beijing, China, 1992.

[2] Huang Wei-Gang and Bai Jin-Xia, "Surges in Substation and Tranformer Windings under Operating Conditions." *High Voltage Engineering*, No. 3, 1990, pp. 9–14.

Manuscript received October 5, 1992.

A. Oguz Soysal and Adam Semlyen: We wish to thank the discussers for their interest in our paper and for their remarks and suggestions. The following are our answers.

To Prof. Huang: The discusser has brought up the interesting and important problem of modal transformations in situations when two or more eigenvalues of the transfer function matrix are nearly equal. The fact that, in the case of exact equality of eigenvalues, the associated eigenvectors are not uniquely defined (so that the particular transformations mentioned by Professor Huang, and many others, are applicable), and only the subspace they span is well determined, results in the strong variation of the transformation matrices in the proximity of such condition. The diagonalization of the original transfer function matrix is therefore not very sensitive on using an exact, frequency dependent transformation matrix and a constant one that is satisfactory over a wider range of frequencies is, we believe, likely to be found.

The discusser's remarks concerning the significance of different degrees of sophistication in modeling are interesting. The estimation procedure described in the paper will ultimately end up in a state equation realization representing the transformer behavior in the time domain. For EMTP applications, the state equations have to be discretized by, say, trapezoidal integration and then the resulting Norton equivalent can be directly used as for any other terminal component.

To Mr. Tsai and Dr. Keyhani: We appreciate the discussers' remarks concerning the possibility of focusing both on the poles and the zeros of the transfer function by using nonlinear least squares estimation, as opposed to the linear least squares approach of the paper. The motivation for using the latter is, of course, its simplicity and robustness (no convergence problems, no initial values needed for the parameters to be estimated). We have also tested cases with noisy data: for a noise level below 3%, the estimation was satisfactory, or else, the order of the approximation had to be increased. By using singular value decomposition, rather than normal equations, we did not have to use the matrix $A^T A$ having the square of the condition number of A itself.

The models given in Table 1 were selected to illustrate the numerical condition of the problem. Transfer functions up to order 20 have also been successfully identified, and one could go higher. Partitioning is particularly appropriate in the case of complex poles close to the imaginary axis. As long as this condition is satisfied, any number of poles can be included in the estimation process. For any segment of partitioning along the frequency axis, the accuracy can be improved by using more poles. Since a complex error criterion is used, the phase error is reduced simultaneously with the error in magnitude. In Figures 2 to 4 we have used only the latter for the purpose of illustration.

To Dr. Morched and Dr. Marti: The suggestion of using orthogonal (Chebyshev) polynomials is, in our opinion, an excellent idea. Clearly, an orthogonal polynomial $P_k(s)$, used instead of the powers s^k in the transfer function (1) of the paper, leads to a much better conditioned problem (6). The reason why orthogonal polynomials have not generally been used for solving the problem of rational approximation is that their fundamental, original role is that of polynomial approximation, i.e., approximation by a power series. This yields

$$\sum a_k P_k(s) \cong H(s) \tag{a}$$

The approximation of $H(s)$ by orthogonal polynomials, shown in (a), is always optimal for a given number of terms, and an additional term will further reduce the error.

Our second remark to this problem is that the use of orthogonal functions for approximation has the advantage that the coefficients a_k in (a) can be determined without solving a set of equations, as $Ax = b$ of (6). Based on the orthogonality of the polynomials $P_k(s)$, we obtain a_k directly by calculating the integral of $H(s)P_k(s)$, much like in the case of a decomposition in a Fourier series.

Since there is no decomposition into orthogonal functions for the rational function approximation

$$\frac{N(s)}{D(s)} \cong H(s)$$

orthogonal polynomials cannot be used in a direct way for the rational polynomial approximation of $H(s)$. However, one may still formulate (starting with an approximate, low order solution) the iterative problem

$$\sum a_{N_k} P_k(s) \cong N(s) = D(s) H(s) \tag{b_1}$$

$$\sum a_{D_k} P_k(s) \cong D(s) = N(s) / H(s) \tag{b_2}$$

where $D(s)$ and $N(s)$ at the right side are assumed to be known from the previous iteration. With this assumption, problem (b) is similar to (a) and can be solved directly for the coefficients of the orthogonal polynomials in $N(s)$ and $D(s)$. It is interesting to note that equations (b) appear to be weighted by $D(s)$ and $N(s)$, respectively, so that when the calculation focuses on coefficients in the numerator, then the inputs near the poles of $H(s)$ are de-emphasized, and vice versa (regarding the denominator and the inputs near the zeros of $H(s)$).

The weighting that emerges from the procedure outlined above increases the emphasis on the object of fitting in the respective iteration steps in (b). In the algorithm used in the paper we did, however, not

apply any additional weighting other than the normal row scaling. The ill-conditioning of the problem has been reduced mainly by column scaling.

The order of the approximation has resulted by starting with a lower order fitting and then gradually increasing the order until the error was sufficiently reduced.

To Dr. Malewski: We thank the discusser for pointing to an important application where terminal measurements on transformers are of interest to utilities, namely, the case of acceptance tests. The time domain measurements, as shown in his plots, are performed at full and reduced voltage levels for the purpose of detecting any dielectric fault. While a transfer function can be identified based on impulse tests, the methodology of the paper does not provide a synthesis of a physical equivalent which could possibly be used for a more detailed diagnosis regarding the internal stresses in the transformer. The state equations that are obtained permit, however, to compute with an Electro-Magnetic Transients Program the voltage at the transformer terminals due to switching or lightning surges.