THE PARAMETRIC INVERSE PROBLEM IN TRANSIENT SCATTERING

D. G. Dudley

Electromagnetics Laboratory
Department of Electrical and Computer Engineering
University of Arizona
Tucson, AZ 85721 USA

INTRODUCTION

Scattering problems in many areas of applied physics are governed by the wave equation. In the most usual situation, we are given the incident wave (input) and the scatterer(s) and attempt, through analytical, experimental, or numerical methods, to produce the scattered waves (output). Such procedures can be carried out in either the frequency domain or the time domain and are categorized under the general heading of "forward problems." In a less usual, but no less important situation, we are given the incident wave (input) and the scattered waves (output) and attempt to find the scatterer(s) that produced the output. In this case, we call the procedures "inverse" problems.

Recently, there has been attention [1,2] to a class of inverse problems with important application in such diverse fields as nondestructive evaluation of materials, target classification, SAR imagery, and geophysical exploration. In this situation, we are given the incident wave (input) and the scattered waves (output), and attempt to find parameters that describe the scatterer(s). We call such procedures "parametric inverse" methods.

In this paper, we describe the parametric inverse problem and discuss the present state of affairs with regard to its solution. We begin with a review of basic results from scattering theory. We next give some system descriptions of transient scattering, defining several problems from the system identification point of view. We follow with a description of a single input, multiple output (SIMO) scattering system and its specialization to single-input, single-output (SISO). We next describe progress in the estimation of parameters for such systems. We conclude with a discussion of the present state of parametric inverse methods and some recommendations for future work. We emphasize that progress in this area depends on better understanding of early time and late time returns from the scatterer(s).

SCATTERING THEORY

Procedures in parametric inverse problems are based on the following result from mathematical scattering theory [3]: Given a scattering
problem governed by the wave equation with Dirichlet boundary conditions, the scattered response to monochromatic plane wave excitation is, with mild mathematical restrictions, a meromorphic function of complex frequency. This result has been extended to the vector wave equation in electromagnetics [4] for scatterers upon whose surface the tangential electric field vanishes. By the Mittag-Leffler theorem [5], the meromorphic property allows the scattered response to be written as a sum over the complex resonances of the scatterer plus an entire function of complex frequency \( s = \sigma + i\omega \). Indeed, if \( v(x,y,z,t) \), abbreviated \( v(r,t) \), is the scattered response as a function of space and time, and \( V(r,s) \) is its Laplace transform,

\[
V(r,s) = \sum_{k=1}^{\infty} \frac{R_k(r)}{s - s_k} + \phi(r,s)
\]

where \( s_k \) are the complex resonances (poles), \( R_k(r) \) are the strengths (residues) of each resonance, and \( \phi(r,s) \) is an entire function. We have assumed in (1) that the poles \( s_k \) are simple. (This property has been postulated for Dirichlet scatterers but has not been proved. There are, however, no known counterexamples.) There are two important characteristics to be noted in (1). First, the residues and the entire function depend on spatial position, while the poles are position independent. Second, the pole series is parametrized by the poles and residues, while the entire function is not parametrized at all.

An inverse Laplace transform of (1) yields the impulse response of the scattering system, viz:

\[
v(r,t) = \sum_{k=1}^{\infty} R_k(r) e^{s_k t} + \phi(r,t)
\]

where \( \phi(r,t) \) is the inverse transform of the entire function. A characteristic of \( \phi \) [6,7] is that it is time-limited. Therefore, for times greater than the time of disappearance of the entire function, the scattering is given completely by the complex exponential series term in (2). Modeling of this complex exponential series has formed the cornerstone in parametric inverse studies to date. Modeling of the entire function remains a subject for present and future study.

**SYSTEM DESCRIPTIONS OF TRANSIENT SCATTERING**

The obtaining of parameters that describe characteristics of the scatterer is based on system descriptions of the scattering process [2]. In the most general case, one has sources at \( m \) spatial locations, \( q \) scatterers, and receivers at \( p \) spatial locations. This case can be modeled as a multiple-input, multiple-output (MIMO) linear system. An initial parametric inverse problem for the MIMO system can be stated as follows:

**Problem I:** Given \( m \) input data records and \( p \) output data records, identify the \( q \) scatterers.

We remark that Problem I is an ambitious undertaking. Indeed, the phase "identify the \( q \) scatterers" is vague and non-specific in parameters.
To obtain a more specific parametric inverse problem, we rewrite (1) as follows:

\[ V(r,s) = \sum_{k=1}^{n} \frac{R_k(r)}{s - s_k} + \epsilon(r,s) \]  

(3)

where \( \epsilon(r,s) \) is an error term given by

\[ \epsilon(r,s) = \phi(r,s) + \sum_{k=n+1}^{\infty} \frac{R_k(r)}{s - s_k} \]  

(4)

Note in (3) that the series "tail" in (1) and the entire function have been combined to form the error term specifically given in (4). For the error to be small, both the tail and the entire function must be small. The tail can often be made small by limiting the data to a restricted frequency range, either by natural limitations on measuring equipment or by intelligent filtering. For times earlier than the disappearance of the entire function, it is not clear that the entire function is small. Indeed, in early time, it often dominates the transient signal. We shall proceed, however, by assuming that the error is small and return to the entire function later in the paper. Provided that \( \epsilon(r,s) \) is small, a second parametric inverse problem can be stated as follows:

Problem II: Given \( m \) input data records and \( p \) output data records, identify the residues \( R_k(r) \) and the poles \( s_k \).

This MIMO system is specific in its parameters. It remains an ambitious undertaking. Progress has been made, however, in several subclasses of the MIMO system, notably, the single-input, multiple-output (SIMO) system and the single-input, single-output (SISO) system, to be described in the following.

THE SIMO SCATTERING SYSTEM

A SIMO model of a scattering system has been proposed by Dudley [2]. In the model, a plane wave is incident on the scattering system from a given direction. In the absence of the scatterer(s), the time history \( u(t) \) of the incident wave is recorded at a reference position \( P(r) \). This time history, adjusted so that \( t = 0 \) corresponds to the initial arrival time of the pulse, is the single input to the system. With the scatterer(s) in place, observations of the total wave are made at \( p \) locations. The time reference of each response is adjusted so that its turn-on time is \( t = 0 \). The incident wave \( u(t) \) is next subtracted from each response to yield the time histories \( y(j)(t) \), \( j = 1, 2, \ldots, p \), corresponding to the components of the scattered wave at different spatial locations. The \( p \) time histories are the multiple outputs from the system. If \( U(s) \) and \( Y(j)(s) \) are the Laplace transforms of the input and output from the \( j \)-th port, respectively, we define the transfer function at the \( j \)-th port by

\[ T(j)(s) = \frac{Y(j)(s)}{U(s)} \]  

(5)

A canonical companion matrix description of the SIMO scattering system can be written in state-space form as follows [2]:

319
\[
\begin{align*}
\dot{x} &= Fx + gu \\
y &= Hx
\end{align*}
\] (6)

where \(x\) is an \(n \times 1\) state vector, \(u\) is the single input, \(y\) is the \(p \times 1\) output vector, and

\[
F = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 & 0 \\
0 & 0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 1 \\
-a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1}
\end{bmatrix}
\] (8)

\[
g = (0, 0, \ldots, 1)^T
\] (9)

\[
H = \begin{bmatrix}
b_0^{(1)} & b_1^{(1)} & \cdots & b_{n-1}^{(1)} \\
b_0^{(2)} & b_1^{(2)} & \cdots & b_{n-1}^{(2)} \\
0 & b_1 & \cdots & b_{n-1} \\
b(p) & b(p) & \cdots & b(p) \\
0 & 1 & \cdots & n-1
\end{bmatrix}
\] (10)

For parametric inversion, let \(y\) be the measured output so that

\[
y = y + e
\] (11)

where \(e\) is an error vector. Therefore,

\[
y = Hx + e
\] (12)

In an SIMO experiment, input-output data \([u, y]\) is substituted into (6) and (12) and the error minimized in some sense by adjustment of the \([a_k, b_k]\). Based on the SIMO model, a third parametric inverse problem can be stated as follows:

**Problem III.** Given the single input data record and the \(p\) output data records, identify the coefficient sequences \([a_k]\) and \([b_k], j = 1, 2, \ldots, p\). Transform the coefficient sequences into the pole sequence \([s_k]\) and the residue sequences \([R_k(j)], j = 1, 2, \ldots, p\).

The transformation from the coefficient sequences into the pole and residue sequences is performed as follows. The Laplace transform of (6) and (7) yields the transfer function

\[
T = \frac{Y}{U} = \frac{H \text{adj}(sI - F) g}{|sI - F|}
\] (13)

where \(I\) is the identity matrix and \(T\) and \(Y\) are column vectors with elements \(T(j)\) and \(Y(j), j = 1, 2, \ldots, p\). If we factor the denominator in (13) into its roots and perform a partial fraction expansion, the result is of the form
where the poles and residues, as required by Problem III, are evident.

THE SISO SCATTERING SYSTEM

For the special case \( p = 1 \), the SIMO scattering system degenerates to SISO. In this case, the state space description is given by

\[
\begin{align*}
\dot{x} &= Fx + gu \\
y &= h^T x
\end{align*}
\]  

(15)

(16)

where \( y \) is the scalar output and \( h \) is an \( n \times 1 \) vector. The equivalence between SISO state space descriptions and differential or difference equation models is well known \[8\]. Indeed, if we time sample (15) and (16) and transform to a difference equation model, the result is

\[
\sum_{j=0}^{n} a_j y(k-j) = \sum_{j=1}^{m} b_j u(k-j) + \varepsilon(k)
\]  

(17)

where \( \varepsilon(k) \) is equation error. In (17), the notation \( \varepsilon(k) \) is short for \( \varepsilon(kT) \), where \( T \) is the temporal sampling interval. The solution to (17) can be given \[9\] in terms of the impulse response \( h(k) \), viz:

\[
h(k) = \sum_{j=1}^{n} R_k s_j e^{jT} + \hat{\varepsilon}(k)
\]  

(18)

where \( \hat{\varepsilon}(k) \) is the error in the exponential model. For the SISO case, represented by (15) and (16), or by (17), a fourth parametric inverse problem can be defined as follows:

Problem IV. Given data on the single input \( u(k) \) and single output \( y(k) \), identify the coefficient sequences \( \{a_k\} \) and \( \{b_k\} \). Transform the coefficient sequences into the pole sequence \( \{s_k\} \) and the residue sequence \( \{R_k\} \).

Problem IV constitutes the major effort to date by researchers attempting to find parametric inverses in transient scattering.

PROGRESS IN THE PARAMETRIC INVERSE PROBLEM

In the previous sections, we have described a structure for examining the parametric inverse problem in transient scattering. The structure is based on canonical descriptions of SIMO systems and their degeneration to SISO systems. In particular, we now consider Problem III and Problem IV described above. Since Problem IV has been the subject of extensive investigation, we give it primary consideration and follow with a short discussion of Problem III.
The solution to Problem IV requires an algorithm for identification of a SISO system. Such an algorithm is NLS [10], an interactive program, based on a generalization of (17). This generalization is contained in the generalized difference equation model of Ljung and Soderstrom [11], given by

\[ A(q^{-1})y(k) = \frac{B(q^{-1})}{F(q^{-1})} u(k) + C(q^{-1}) d(k) e(k) \]  \hspace{1cm} (19)

where

\[ q^{-1} u(k) = u(k-1) \]  \hspace{1cm} (20)

and \( A, B, C, D, F \) are polynomial stepping operators with, typically,

\[ A(q^{-1}) = \sum_{n=0}^{N} a_n q^{-n} \]  \hspace{1cm} (21)

The equation error model given by (17) is a special case of (19), obtained by setting \( C = D = F = 1 \), to give

\[ Ay(k) = Bu(k) + e(k) \]  \hspace{1cm} (22)

The coefficients \( A \) and \( B \) are identified by minimizing the Euclidean norm of the error \( e(k) \). This identification is a problem in linear least squares [9]. The coefficients so identified are used as starting values to solve an output error model, given by setting \( A = C = D = 1 \), to give

\[ y(k) = \frac{B}{F} u(k) + e(k) \]  \hspace{1cm} (23)

The solution of the output error model is done in NLS by a nonlinear optimizer that conducts a search for the global minimum of the Euclidean norm of \( e(k) \). For details, the interested reader is referred to [10,12]. The solution yields the coefficients in (23). The poles and residues are then found [9] by taking the z-transform of (23), factoring the characteristic denominator polynomials, performing a partial fraction expansion, and finally applying the transformation

\[ s = \frac{\ln z}{T} \]  \hspace{1cm} (24)

The procedure described above, leading to the solution to Problem IV, has been the subject of investigation with both simulated and real data in the study of electromagnetic transients. In the simulated data case, Goodman and Dudley [12] have concentrated on backscatter from a dielectric slab, a geometry studied in depth in a doctoral dissertation by Nabulsi [13]. The principal advantage of the slab as a test case is that the entire function in (1) can be removed analytically from the transient scattered signal. Therefore, the error term in (4) consists of only the tail of the pole series. An additional feature of the slab is that its resonances are prominent and well-separated in the frequency domain. Goodman and Dudley find that, with judicious low-pass filtering,
the tail of the pole series can be made small enough so that the lower order resonances (lowest frequencies) are clearly and accurately identified. In particular, they adjust the low-pass filter so that there are seven resonances in the pass band. In the absence of noise (except for computer roundoff), they find that NLS reproduces the seven resonances to a high degree of accuracy.

This positive result deteriorates in the presence of noise. Goodman and Dudley use the signal processing algorithm SIG [14] to add one-percent, five-percent, and ten-percent noise to the scattered signal. (By "one-percent" noise, they mean that the noise standard deviation is 0.01 times the peak value.) For one-percent noise, the damping term in the complex poles begins to deviate from the theoretical result after the third pole-pair. This result is consistent with the discussion in [9]. For five-percent noise there is good agreement for only the first two pole pairs. For ten-percent noise, only the first pole pair has good agreement with theory.

In a study with real data, Dudley and Goodman [15] report results in electromagnetic transient scattering from a flat conducting rectangular plate, 30 cm high and 60 cm wide. The experiment was performed on the Lawrence Livermore National Laboratory transient range. In this case, the scattered field contains both the pole series and the entire function. It is, therefore, to be anticipated that it will be more difficult to perform an identification by the guidelines of Problem IV, where the entire function is treated as an error in the output error model. Dudley and Goodman assess the value of the identification as a parametric inverse by comparing the results with the theoretical natural resonances for the plate obtained by Pearson [16]. They find good agreement between the theoretical and identified lowest order (lowest frequency) pole pairs. Physically, this resonance corresponds to the major body resonance of the scatterer. This agreement is important since it gives an estimate of the physical size of the object.

Note that both of the results reported above are for backscatter. Indeed, the SISO model has the limitation that it fails to utilize the additional degree of freedom provided by measurements at different spatial locations. This limitation is overcome by the SIMO model. Although Dudley has discussed the model [2], he as yet has given no tests with data. Goodman and Dudley [12], however, have reported an extension to NLS that includes the SIMO case and results with both simulated and experimental data are planned.

DISCUSSION

We have discussed the parametric inverse problem in transient scattering and have reported on progress to date in its solution. We have specified several problems and have noted that the principal work to date has been in attempts to solve Problem IV. The value of results in Problem IV as a parametric inverse, however, is limited by three factors. First, in most scattering problems of practical interest, the radiation damping causes a rapid decay in the time history of the scattered field. This effect leads to a severe limitation on the number of data points available in an identification experiment. Unfortunately, this difficulty seems to be fundamental. Second, Problem IV is SISO and, therefore, ignores the possibility of multiple angle observations. Generalizations of NLS to SIMO should make possible the study of this additional degree of freedom. Third, the entire function has not been included in any presently available models. This omission is unfortunate since a large portion of the early time response is often contained therein.
Dudley [2] has suggested augmenting the state-space description of the SIMO model with an early time descriptor. Indeed, let \( \phi^{(j)}(s) \) be the entire function contribution to the transfer function \( T^{(j)}(s), j = 1, 2, \ldots, p \), and write

\[
\phi^{(j)}(s) = \sum_{k=1}^{q} c_k^{(j)} s^{-k}
\]  

We remark that the multiple-order pole at \( s = 0 \) in (25) is inconsistent with the entire function characteristic. Our interest, however, is at high frequency. Equation (25) can be incorporated into the SIMO model by augmenting the vectors and matrices in the state space description in (6) and (7). The model is appealing because the terms are precisely the form as asymptotic expansions [17] of the scattered response at high frequencies (descending powers of frequency). These terms contain information on radius of curvature and reflection coefficient, parameters that are as yet missing in the parametric inverse problem. The utility of this formulation, however, must await the outcome of present and future studies.

REFERENCES


DISCUSSION

Mr. Brian DeFacio (University of Missouri): Is is possible theoretically or experimentally to learn about the entire function by picking cases where the residues are small? Can you get a residue which is numerically small?

Mr. Dudley: The question is: Is it possible to learn about the entire function by picking places where the residue is small? The answer is that for practical objects there is a sequence of poles. When you monitor the scattered response at different locations around the object some of the residues decrease and others increase. It is difficult to think of a case where they would all be numerically small at the same time.

Mr. Johnson (National Bureau of Standards): We have been interested in problems very similar to the ones you have been discussing. We also have non-linear optimization problems that we are developing at NBS. We have techniques that I have developed for doing what you would call the generalized cloning method by least squares and I would be glad to show them to you. It is probably the most powerful method around. We can pull out eight to ten poles using the technique. You mentioned the problem of obtaining the roots of the characteristic polynomial. The key for this is an algorithm based on a new polynomial root solver I have developed.

Mr. Dudley: A lot different than Jenkins and Traub?

Mr. Johnson: Infinitely superior.

Mr. Dudley: I would like to talk with you.

Mr. A. J. Devaney (Devaney Associates): Don, is there any relationship between defining these complex poles and the work being done on measuring time of arrival? I think of a pulse arriving, especially in the case of a plate and I am either looking at the resonances or I am looking at the first arrival back, the second, and so on. Parameter estimation can be done on these arrival times to obtain the parameters and you are doing it in the frequency domain. Is there a relationship between the two?

Mr. Dudley: Consider the slab. In that case there are only two parameters of interest: the dielectric constant and the thickness. These parameters can be obtained more directly from the various times of arrival than from the complex resonances. We use the slab to simply indicate the complex resonance estimation. In most of the more complicated problems, for instance, for the sphere . . .

Mr. Devaney: It is not so clear cut.
Mr. Dudley: For the sphere, it is still fairly clear cut. What you see is the first return and then the first creeping wave. Again you are seeing discrete events that are clearer in terms of resonances. Resonance identification becomes the clear cut way only when resonances are highly separated and have a very high Q. In such cases complex resonance identification makes a lot of sense.

Mr. Devaney: I guess what I am asking is this: Is there a one-to-one mapping between resonances and arrival times?

Mr. Dudley: There is for problems with no radiation damping, such as the slab. In general, for structures with radiation damping, where the shape of the return varies with each arrival, the picture is not so simple.