SIAM REVIEW
Vol. 29, No. 3, September 1987

© 1987 Society for Industrial and Applied Mathematics

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

INVERSE SCATTERING FOR DISCRETE TRANSMISSION-LINE MODELS*

A. M. BRUCKSTEIN†§ AND T. KAILATH†‡

Abstract. This paper presents several methods for the identification of the impedance and reflection coefficient profile of a nonuniform, discrete transmission-line from its response to a given forcing function. This problem is the prototype for a wealth of one-dimensional *inverse scattering* problems arising in various fields. A unified, straightforward approach is presented, providing all known inversion procedures and some novel ones. The derivations readily follow from causality of signal propagation on the transmission-line. It is shown that a direct exploitation of the signal propagation model leads to recursive, computationally efficient and reliable inversion algorithms. In particular, the relationships between *layer-peeling* (Schur type, difference equation) and *layer-adjoining* (Levinson type, integral equation) algorithms are clearly displayed.

Key words. fast algorithms, Schur recursions, layer-peeling methods, Gelfand–Levitan equations, lossy media

AMS(MOS) subject classifications. 15-02, 15A06, 15A90, 45A05, 45E10, 81F99, 93E11

A mathematical theory is not to be considered complete until you have made it so clear that you can explain it to the first man whom you meet on the street.

David Hilbert (quoting an old French Mathematician),
in *Mathematical Problems*, AMS Bulletin 8, 1901–02.

1. Introduction. Motivated by numerous applications, physicists, electrical engineers, geophysicists and mathematicians have studied a wide range of inverse scattering problems. It turns out that inverse spectral problems for Schrödinger operators and vibrating strings, algorithms for transmission-line synthesis, determination of the vocal tract area function in speech research, identification of the acoustic impedance or conductivity profiles in layered-earth models, the design of digital filters in cascade form and the derivation of fast lattice-form linear least-squares prediction algorithms, can all be interpreted as sharing a common mathematical foundation. They all require a procedure that identifies from the so-called *scattering data* (which are "excitation-response" pairs of time-functions measured at the boundary) the depth-dependent parameters of a one-dimensional, highly structured signal propagation model, whose prototype is, for electrical engineers, a lossless nonuniform transmission-line.

A famous paper of Gelfand and Levitan, published in 1955, reduced the solution of the (Schrödinger) inverse scattering problem to the solution of a parametrized set of linear integral equations. Subsequent work of Marchenko, Krein and others, further strengthened the belief that inverse scattering is equivalent to solving linear integral or, in the discrete case, linear matrix equations (see e.g. [1]–[7]).

However, in some quite early work, mostly in geophysics, a different approach emerged. This approach, more directly concerned with a local analysis of acoustic

* Received by the editors September 1, 1984; accepted for publication (in revised form) November 25, 1985. This work was supported in part by the U.S. Army Research Office under contract DAAG29-79-C-0215 and by the Air Force Office of Scientific Research under contract AF49-620-79-C-0058.

† Information Systems Laboratory, Stanford University, Stanford, California 94305.

‡ This author gratefully acknowledges the support provided by an Erna and Jakob Michael Visiting Chair in Theoretical Mathematics at the Weizmann Institute of Science, Rehovot, Israel, during Spring Quarter, 1984.

§ Present address, Faculty of Electrical Engineering, TECHNION, Haifa, Israel.

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

wave propagation, provided recursive methods, called dynamic deconvolution, or downward continuation, algorithms (see e.g. [8]–[13]), for the identification of layered-earth models which are the geophysical counterpart of nonuniform transmission lines. Though by now fairly well known in some of the geophysics and engineering literature, such direct methods were not always regarded as viable, since the effect of noise in the scattering data was intuitively expected to be disastrous. Most researchers therefore returned to integral or matrix-equation-based approaches with the hope that such methods would provide some averaging of the noise and thus have better numerical behavior. Though this expectation was never justified, since both classes of methods produce the same results assuming infinite precision computations and also have similar numerical stability properties (see [14]–[15]), almost all effort in "inverse scattering" research has been focused on exploration of linear equation formulations, providing alternative derivations, interpretations, etc. (see e.g. [16]–[21]). Only recently was the balance somewhat restored as more work has been devoted to a careful analysis of direct, differential methods (see e.g. [22]–[27]). In so doing, it has been noticed that, though both direct and indirect methods are numerically stable, the very structure, and the simplicity of derivation, of the direct methods indicates intelligent "error control" techniques that allow more effective medium identification in the presence of noise in the scattering data (see [28], [29]).

The main goal of this paper is to use a discretized form of the nonuniform transmission-line equations as the vehicle for simple, physically based derivations and comparisons of the direct (downward continuation type) and indirect (or linear equations based) methods. We briefly indicate some of the new insights and new results of this paper here.

As a first example we mention that in our approach (see §4) it becomes clear that the linear equations of the inverse scattering problem are direct consequences of (only) the causality and symmetry properties of the transmission-line model. In fact we obtain an apparently new, general set of linear equations for solving the inverse problem. This set of equations is based on arbitrary scattering data (i.e. the response to an arbitrary excitation function) and we show that the discrete analogues of the classical equations of Gelfand–Levitan, Marchenko and Krein are obtained by making special choices of excitation-response pairs.

More centrally, our theme is that formulation via the discrete transmission-line model shows (see §3) that a simple layer-peeling inverse scattering procedure arises more directly than do the linear equations; this procedure only invokes causality of signal propagation and a simple rule for recursively computing the discrete waveforms at increasing depths in the scattering medium. For lossless media, the layer-peeling method is similar to the dynamical deconvolution (downward continuation) method discovered by geophysicists (see [9], [10], [29]). Our simple approach highlights the fact that the layer-peeling method also applies not only to lossless scattering media but to several more general models, with nonsymmetric and even nonlinear interactions arising in applications such as the minimal partial realization problem of system theory [30], the development of efficient decoding algorithms for error correcting codes [31], and in solving certain nonlinear convolution-type integral equations arising in biological modeling problems (see [32]).

Moreover it is fascinating to see (§3) that, in the lossless case, one version of the direct inversion algorithm, when expressed analytically via the so-called z-transform (or generating function) notation for discrete signal analysis, turns out to be the same as a recursive procedure devised by I. Schur in 1917 (see [33], [34]), for testing if a power series in z^{-1} is bounded outside the unit disc in the complex plane. This connection of the Schur algorithm to inverse scattering was perhaps first explicitly

made by P. Dewilde (ca. 1977), who had encountered the Schur algorithm in studies of cascade synthesis procedures for stable digital filters (see [35]–[38] and also [39]). As mentioned above, geophysicists had unknowingly rederived the Schur algorithm in their development of the downward continuation (dynamic deconvolution) method [9]–[13].

The direct inversion methods have a complexity of $O(N^2)$, i.e., proportional to N^2 computations are needed to identify the medium up to depth N. The solution via linear equations, when those are solved in a straightforward way by, say, Gaussian elimination, would require $O(N^3)$ computations. The computational efficiency of the direct methods is due to the fact that these fully exploit the assumed structure of the scattering medium. Linear-equations-based methods can be made to yield computationally efficient algorithms if the special structure of their coefficient matrices is cleverly taken into account; this was done, for example, in [17], [18], [40] for the important special equations of Gelfand–Levitan, Marchenko and Krein. We shall see that the understanding of how the medium structure leads to the general linear equations for inverse scattering also provides the appropriate (general) fast algorithms including as particular cases the algorithms derived in the literature for the Krein or Marchenko and Gelfand–Levitan equations.

In brief, we show in this paper that the transmission-line model clearly brings out the reasons for various differences between the direct (*layer-peeling*) and linear equations (which, as we shall see, may be called *layer-adjoining*) methods and in fact readily suggests generalizations of various existing results. Furthermore, we can easily recognize the possibility of implementing layer-peeling methods with parallel computation, requiring $O(N)$ time with an array of N processors, or with doubling techniques that require $O(N \log^2 N)$ time on a single processor.

We should emphasize that the transmission-line model on which the arguments are based is quite general: by appropriate transformations, a variety of other inverse problems can be recast in this form, e.g., the discrete versions of the vibrating string problem [41], the inverse Schrödinger problem [5]–[7], and the acoustic vocal tract problem [42]. The results for the corresponding continuous problems can be obtained by a limiting procedure, or by following similar, though somewhat less elementary, "propagation of singularities" arguments [20], [25], [27].

However we should note that the problems solvable by the approaches discussed in this paper are inherently one-dimensional. Inverse scattering in more than one spatial dimension is as yet an unsolved problem in all but a few particular cases with special symmetries that effectively reduce the dimensionality. It is however a problem of much interest in geophysics and materials science, concerned with nondestructive testing and in medical imaging with ultrasound waves, and therefore a lot of research is dedicated to it. In stratified media with general geometries, even the direct scattering problem, i.e. the analysis of wave propagation, is in many cases too difficult to tackle. Solving these problems requires in most cases complicated ray-tracing methods and also accounting for the possibility of having guided waves along internal layers in which propagation speed is higher. As far as we know, important problems such as what is the scattering data that has to be gathered in order to make a high-dimensional inversion problem well-posed are not completely solved today. It is, however, possible that a layer-peeling idea combined with some ray-tracing method would lead to a successful inversion algorithm for certain particular cases of the multi-dimensional inverse scattering problem. Some work in this direction is described in [43].

This paper is organized as follows. A preliminary discussion of basic facts on signal propagation on lossless, discrete transmission lines is given in §2. Several direct algorithms for inverse scattering and their relationships to the work of Schur are then

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

derived in §3. In §4, we use the transmission-line picture to develop linear equation approaches to the inverse problem. Several remarks on related results in factorization and estimation theory conclude our discussion of inverse scattering methods.

2. Discrete lossless transmission lines. Let us review some basic facts about nonuniform and lossless transmission-line models. The propagation of signals on a lossless transmission line can be described by a set of symmetrized, so-called "telegrapher's equations" [25], [44], [45],

$$\begin{aligned} \frac{\partial}{\partial x} v(x,t) &= -Z_x \frac{\partial}{\partial t} i(x,t), \\ \frac{\partial}{\partial x} i(x,t) &= -Z_x^{-1} \frac{\partial}{\partial t} v(x,t) \end{aligned} \tag{2.1}$$

where $v(x, t)$ is the voltage at point x on the line, at time t, $i(x, t)$ is the current at point x at time t, and Z_x is the (local) impedance at the point x. By assumption Z_x is strictly positive and finite, and, without loss of generality, we shall set $Z_0 = 1$. Normalizing the voltage and current variables to

$$V(x,t) = v(x,t)(Z_x)^{-1/2} \quad \text{and} \quad I(x,t) = i(x,t)(Z_x)^{1/2} \tag{2.2}$$

at each point on the line, we get for the $\{V(x, t), I(x, t)\}$ pair of variables the equation

$$\frac{\partial}{\partial x}\begin{bmatrix} V(x,t) \\ I(x,t) \end{bmatrix} = \begin{bmatrix} -k_x & -\partial/\partial t \\ -\partial/\partial t & k_x \end{bmatrix}\begin{bmatrix} V(x,t) \\ I(x,t) \end{bmatrix} \tag{2.3}$$

where k_x is the so-called local "reflectivity" parameter given by

$$k_x = \frac{1}{2}\frac{d}{dx} \ln Z_x. \tag{2.4}$$

It is useful to derive a modified set of equations relating the so-called left- and right-propagating *wave variables* defined via the transformations

$$W_R(x,t) = \frac{V(x,t) + I(x,t)}{2} \quad \text{and} \quad W_L(x,t) = \frac{V(x,t) - I(x,t)}{2}. \tag{2.5}$$

For these quantities the evolution equations, obtained by combining (2.3) and (2.5), are

$$\frac{\partial}{\partial x}\begin{bmatrix} W_R(x,t) \\ W_L(x,t) \end{bmatrix} = \begin{bmatrix} -\partial/\partial t & -k_x \\ -k_x & \partial/\partial t \end{bmatrix}\begin{bmatrix} W_R(x,t) \\ W_L(x,t) \end{bmatrix}. \tag{2.6}$$

This set of equations is interpreted as describing the propagation of waves to the left and right, the intensity of their local interaction being parametrized by the reflectivity k_x. Indeed, suppose that over a portion of the line, say $[x_a, x_b]$, the local impedance is constant. Then, by (2.4), the local reflectivity is zero and the above equations provide

$$\begin{aligned} \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right) W_R(x,t) &= 0, \qquad x_a < x < x_b, \\ \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\right) W_L(x,t) &= 0, \qquad x_a < x < x_b, \end{aligned} \tag{2.7}$$

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with general solutions of the form: $W_R(x, t) = W_R(x - t)$ and $W_L(x, t) = W_L(x + t)$, describing noninteracting (left and right) propagating waves. Therefore, any portion of a transmission line with constant local impedance acts on the wave variables as a pure time-delay and time-advance operator. When k_x is nonzero, the propagating waves do interact, part of each wave being "backscattered" and added to the one propagating in the opposite direction.

2.1. Discrete wave propagation equations. With this background, let us consider a discrete transmission line with sections of constant impedance over intervals of length $\frac{1}{2}$ (see Fig. 1). We shall define

$$Z_x = Z_{n-1}, \qquad x \in [(n-1)/2, n/2) \tag{2.8}$$

as the impedance of the nth section, and recall our assumption that $Z_0 = 1$. Because of the constancy of the local impedance over a section, we can write (see (2.7))

$$\begin{bmatrix} W_R(x,t) \\ W_L(x,t) \end{bmatrix}_{x=(n/2)-} = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta^{-1} \end{bmatrix} \begin{bmatrix} W_R(x,t) \\ W_L(x,t) \end{bmatrix}_{x=[(n-1)/2]+} \tag{2.9}$$

where Δ is a "time-delay" operator defined by

$$\Delta f(t) = f(t - \tfrac{1}{2}) \tag{2.10a}$$

and correspondingly Δ^{-1} is a "time-advance" operator

$$\Delta^{-1} f(t) = f(t + \tfrac{1}{2}). \tag{2.10b}$$

Note that the time-delay operator is physically meaningful for right-going waves, while the time-advance operator is so for left-going waves.

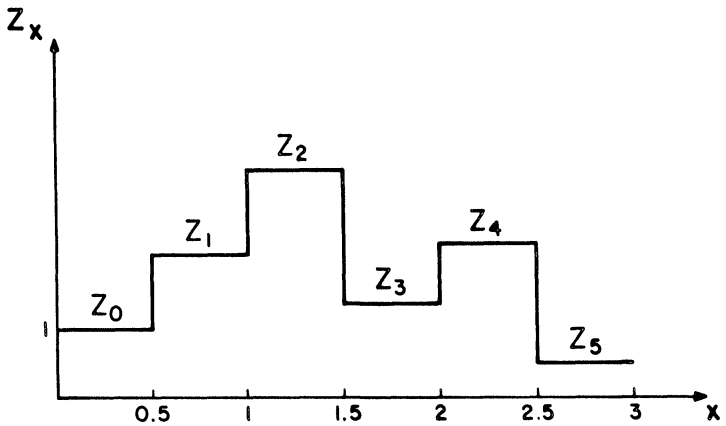

FIG. 1. *Typical impedance profile for a discrete transmission line.*

Wave interactions, reflections and transmissions, will occur at the points where there are changes in the local impedance. To compute these interactions, we return to the original unnormalized voltage and current variables $\{v, i\}$, which by (2.1) have to be continuous at all points of the line, i.e., we have

$$\begin{bmatrix} v(x,t) \\ i(x,t) \end{bmatrix}_{x=(n/2)+} = \begin{bmatrix} v(x,t) \\ i(x,t) \end{bmatrix}_{x=(n/2)-}. \tag{2.11}$$

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Using (2.2) and (2.5), we see that

$$\begin{bmatrix} W_R(x,t) \\ W_L(x,t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} Z_x^{-1/2} & Z_x^{1/2} \\ Z_x^{-1/2} & -Z_x^{1/2} \end{bmatrix}\begin{bmatrix} v(x,t) \\ i(x,t) \end{bmatrix} = \Xi(Z_x)\begin{bmatrix} v(x,t) \\ i(x,t) \end{bmatrix}, \tag{2.12}$$

say. Therefore we obtain, combining (2.11) and (2.12),

$$\begin{bmatrix} W_R \\ W_L \end{bmatrix}_{x=(n/2)+} = \Xi(Z_n)\Xi^{-1}(Z_{n-1})\begin{bmatrix} W_R \\ W_L \end{bmatrix}_{x=(n/2)-} = \Theta_n\begin{bmatrix} W_R \\ W_L \end{bmatrix}_{x=(n/2)-}. \tag{2.13}$$

Note that we wrote W_R instead of $W_R(x,t)$, and so on, the independent variables being clearly understood from the context. Now referring back to (2.9), we can write a complete discrete wave-evolution equation,

$$\begin{bmatrix} W_R \\ W_L \end{bmatrix}_{x=(n/2)+} = \Theta_n\begin{bmatrix} \Delta & 0 \\ 0 & \Delta^{-1} \end{bmatrix}\begin{bmatrix} W_R \\ W_L \end{bmatrix}_{x=[(n-1)/2]+} \tag{2.14}$$

where

$$\Theta_n = \Xi(Z_n)\Xi^{-1}(Z_{n-1}) = \frac{Z_n + Z_{n-1}}{2(Z_n Z_{n-1})^{-1/2}}\begin{bmatrix} 1 & -\dfrac{Z_n - Z_{n-1}}{Z_n + Z_{n-1}} \\ -\dfrac{Z_n - Z_{n-1}}{Z_n + Z_{n-1}} & 1 \end{bmatrix}. \tag{2.15}$$

Define the nth local "reflection coefficient" as

$$k_n = (Z_n - Z_{n-1})/(Z_n + Z_{n-1}). \tag{2.16}$$

Since $Z_n > 0$ for all n, it follows that

$$|k_n| \leqq 1. \tag{2.17}$$

The introduction of the reflection coefficient sequence $\{k_n\}$ allows rewriting of the matrix Θ_n in the form

$$\Theta_n = (1 - k_n^2)^{-1/2}\begin{bmatrix} 1 & -k_n \\ -k_n & 1 \end{bmatrix} \tag{2.18}$$

and it is easy to check that such matrices are J-orthogonal, i.e.

$$\Theta_n J \Theta_n^T = J, \qquad J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.19}$$

where superscript T denotes transpose. This relationship, expressing energy conservation, can also be derived in a more physical way, as will be seen below.

For simplicity of notation we shall define

$$\begin{bmatrix} W_R(n,t) \\ W_L(n,t) \end{bmatrix} = \begin{bmatrix} W_R \\ W_L \end{bmatrix}_{x=(n/2)+} \tag{2.20}$$

so that the relation (2.14) will be written

$$\begin{bmatrix} W_R(n,t) \\ W_L(n,t) \end{bmatrix} = \Theta_n\begin{bmatrix} \Delta & 0 \\ 0 & \Delta^{-1} \end{bmatrix}\begin{bmatrix} W_R(n-1,t) \\ W_L(n-1,t) \end{bmatrix}. \tag{2.21}$$

A pictorial signal-flow graph "representation" of the wave propagation implied by (2.21) is shown in Fig. 2(a). (See Appendix A for a review of signal-flow graph representations.) We use the quotation marks to emphasize that, due to the presence

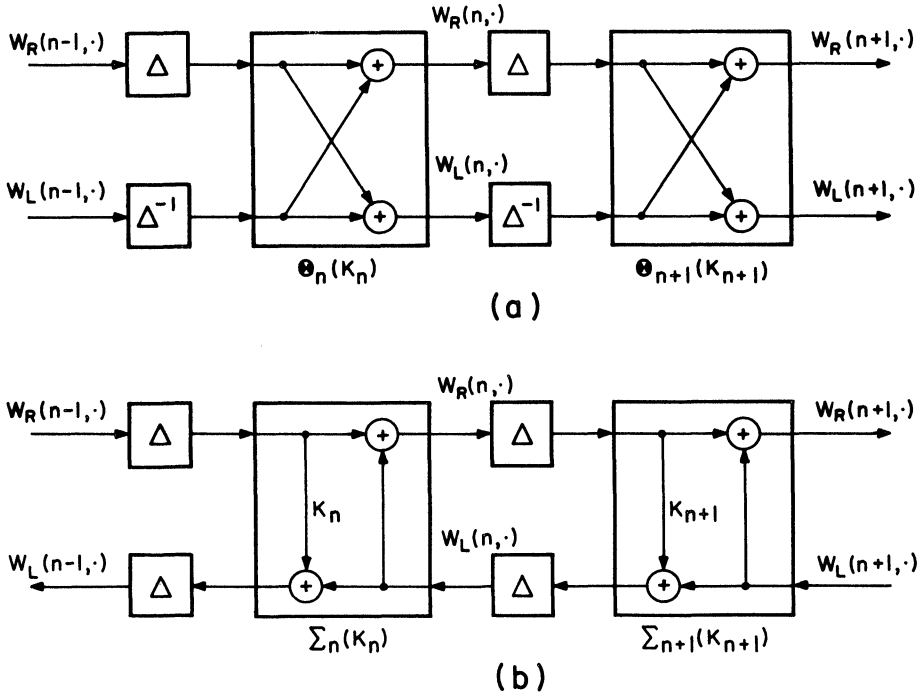

FIG. 2. *The wave propagation model* (a) *transmission diagram* (b) *causal scattering picture.*

of the advance operator Δ^{-1}, this scheme does not correspond to a causal/physical signal flow.

A physically meaningful representation can be obtained by reversing the direction of signal flow for the wave W_L, with the effect that the "time-advance" operator in (2.21) becomes a causal, "time-delay" operator Δ acting on the left propagating wave. Of course, the wave interaction matrix Θ will have to be modified so as to preserve the algebraic relationships between W_R and W_L at each point of the line. The proper modification can be obtained by standard flow-graph manipulation rules or by simple algebraic rearrangement (see e.g. [46] or Appendix A). The result is that (2.21) can be rewritten as a *causal wave scattering* equation

$$\begin{bmatrix} W_R(n,t) \\ W_L(n-1,t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \Delta \end{bmatrix} \Sigma_n \begin{bmatrix} \Delta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} W_R(n-1,t) \\ W_L(n,t) \end{bmatrix} \tag{2.22}$$

where the scattering or interaction-gain matrix Σ is given by

$$\Sigma_n = \begin{bmatrix} (1-k_n^2)^{1/2} & -k_n \\ k_n & (1-k_n^2)^{1/2} \end{bmatrix}. \tag{2.23}$$

A sequence of such causal wave scattering relationships is graphically represented by the now physically meaningful discrete transmission-line model shown in Fig. 2(b).

The matrix Σ_n is called the *scattering matrix* of the nth section of the line since it describes the physical interaction between the incident waves $\{W_R(n,t), W_L(n+1,t)\}$ and the reflected waves $\{W_R(n+1,t), W_L(n,t)\}$ at that section. It is easy to check that Σ_n is unitary, i.e.,

$$\Sigma_n^T \Sigma_n = I = \Sigma_n \Sigma_n^T, \tag{2.24}$$

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

which corresponds to the physical property of energy conservation or losslessness, i.e., ($\|f(\cdot)\|$ denoting l_2 norm)

$$\|W_R(n-1,\cdot)\|^2+\|W_L(n,\cdot)\|^2=\|W_R(n,\cdot)\|^2+\|W_L(n-1,\cdot)\|^2, \tag{2.25}$$

or *incident energy equals reflected energy*. Note also that k_n can indeed be interpreted as a local reflection coefficient because from (2.14) we have

$$k_n=\frac{W_L(n-1,t)}{W_R(n-1,t-1)} \quad \text{if } W_L(n,\tau)=0 \text{ for } \tau<t. \tag{2.26}$$

An immediate consequence of the energy conservation relation (2.25) is that

$$\|W_R(n,\cdot)\|^2-\|W_L(n,\cdot)\|^2=\|W_R(n-1,\cdot)\|^2-\|W_L(n-1,\cdot)\|^2, \tag{2.27}$$

which explains why the matrix Θ_n obeys the previously noted identity

$$\Theta_n J\Theta_n^T=J, \qquad J=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}.$$

The matrices $\{\Theta_n\}$ are often called *chain scattering* or *transfer* matrices. The transfer representation of a wave-scattering process, although not always corresponding to a physical signal flow, is very useful because the natural cascade composition rule for transfer matrices is simply the usual matrix multiplication. For scattering matrices $\{\Sigma_n\}$, the cascade composition requires a more involved computation known as Redheffer's star-product rule [46].

Among the more significant consequences of energy conservation relations are results on fast algorithms for the triangular factorization of a class of matrices that can be regarded as natural generalizations of the much-studied Toeplitz matrices, see [47]. We shall not pursue these interesting connections here (see also [48]), but shall turn instead to evolution equations for the voltage and current variables.

2.2. Voltage-current evolution equations. Though the physical evolution of the wave variables $\{W_R, W_L\}$ has attracted most attention so far, it should be noted that there are similar evolution equations for the voltage and current variables $\{v, i\}$, which are equally useful in inverse scattering (as we shall see in §3.1).

Since the $\{v, i\}$ variables are continuous along the whole line, it suffices to consider their propagation *within* a single section. The propagation equations can be obtained by combining (2.9) and (2.12), and we obtain

$$\begin{bmatrix}v\\ i\end{bmatrix}_{x=n/2}=\Xi^{-1}(Z_{n-1})\begin{bmatrix}\Delta & 0\\ 0 & \Delta^{-1}\end{bmatrix}\Xi(Z_{n-1})\begin{bmatrix}v\\ i\end{bmatrix}_{x=(n-1)/2} \tag{2.28}$$

where $\Xi(Z)$ is from (2.12)

$$\Xi(Z)=\frac{1}{2}\begin{bmatrix}Z^{-1/2} & Z^{1/2}\\ Z^{-1/2} & -Z^{1/2}\end{bmatrix}. \tag{2.29a}$$

This matrix has the scattering representation

$$\Xi_\Sigma(Z)=\begin{bmatrix}2Z^{-1/2} & -1\\ Z^{-1} & -Z^{-1/2}\end{bmatrix}, \tag{2.29b}$$

and therefore the voltage-current propagation can be described pictorially as in Fig. 3, either as a noncausal signal transfer flow-graph (3(a)) or as the corresponding, more physical, scattering diagram (3(b)).

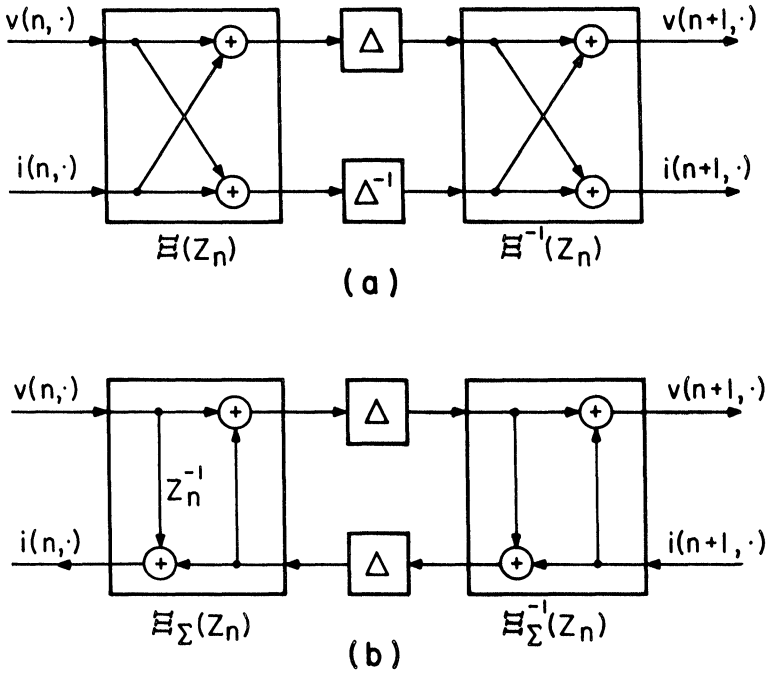

FIG. 3. *The current-voltage propagation model* (a) *noncausal signal transfer* (b) *scattering diagram.*

For completeness, we also note here that from (2.16) we can express the local impedance of the discrete line in terms of the local reflection coefficient sequence as

$$Z_n = \frac{1+k_n}{1-k_n} Z_{n-1} = \cdots = \prod_1^n \frac{1+k_i}{1-k_i} \tag{2.30}$$

where we again recall our assumption that $Z_0 = 1$.

Gain-delay characterizations. In summary, note that the various formulations given above provide a so-called *gain-delay* characterization of the transmission-line. The line is seen to be a layered, or cascade, system composed of static "gain" sections separated by delays. Each gain section is specified completely by a wave interaction matrix that depends on a local parameter, which can either be the reflection coefficient or the local impedance. We shall see that this characterization immediately suggests a direct, time-domain recursive procedure for solving the inverse problem for lossless lines. Moreover, the same inversion method will be applicable to more general wave propagation models having similar gain-delay structure.

3. Layer-peeling algorithms for inverse scattering. The inverse scattering problem for transmission-lines is to determine the line given the input and response functions, $W_R(0, t)$ and $W_L(0, t)$, under the assumption that the line was initially quiescent. This assumption means that the response recorded is entirely due to the causal propagation of the probing signal on the line.

There are many ways to attack this problem. The most widely known are methods using special choices of input sequences, based on which the inversion problem is shown to be equivalent to the solution of sets of linear equations, discrete analogues of certain well-known integral equations of continuous inverse scattering, associated

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with the names of Gelfand and Levitan [1], [4], Krein [3], Marchenko [2], Gopinath and Sondhi [45] (see also [17], [20], [25], [27]).

The linear matrix equations required to determine the line up to length N are of increasing size, up to $N \times N$, so that their solution would generally require $O(N^3)$ elementary operations. However, the special properties of the transmission-line impose further structure on the linear equations, e.g., making their coefficient matrices Toeplitz (Krein, Gopinath–Sondhi), Toeplitz plus Hankel (Gelfand–Levitan) or Hankel (Marchenko, Burridge). These special structures can be exploited to reduce the computational burden by an order of magnitude to $O(N^2)$, leading to so-called fast inverse scattering algorithms (see e.g. [17], [18], [40]). It turns out, however, that this is a rather indirect route to fast inverse scattering algorithms. The special properties of transmission-lines, analyzed in detail in §2, directly suggest a family of fast inverse scattering algorithms, which were apparently first discovered by geophysicists and called "dynamic deconvolution" algorithms (see [8]–[10]).

In this section, we develop several direct inversion algorithms in a form that immediately suggests extensions to more general inverse scattering problems. First we shall show that the gain-delay transmission-line structure, described in §2, can be used to obtain a general direct inversion algorithm. This algorithm uses the scattering data as given, and there is no need for a "deconvolution" step to put the data in any special form. The intimate relationship of the direct inversion procedure to certain results of Schur [33] is then discussed in some detail. Finally we show how these ideas can readily be extended to more general inverse problems for lossy and even a class of nonlinear scattering media.

3.1. "Layer-peeling" algorithms for lossless lines. We start with a discrete transmission-line, i.e., one with a piecewise constant impedance profile. The line is assumed to be quiescent at $t = 0$, i.e., $v(x, 0) = 0 = i(x, 0)$, when we apply a current input at $x = 0$ and measure the voltage response at that point, say $v(0, t)$. The problem is to recover the parameters of the transmission-line, e.g., the local reflection coefficients or the local impedances at all depths x in this scattering medium. Of course, we can easily go from one parametrization to the other by using the relations (2.16) and (2.30); however it turns out to be just as easy, and more advantageous in practice, to determine each of them directly.

Recovery of the local reflection coefficients. The first step is to transform the data into the wave variables at $x = 0$ (cf. (2.5)). We get the equivalent scattering data

$$W_R(0,t) = [V(0,t) + I(0,t)]/2 = [v(0,t) + i(0,t)]/2, \tag{3.1}$$
$$W_L(0,t) = [V(0,t) - I(0,t)]/2 = [v(0,t) + i(0,t)]/2,$$

since, by assumption, $Z_0 = 1$.

Now from (2.22) (see Fig. 2(b)) and because the line is assumed to be at rest at $t = 0$, there will be no left propagating wave response from the line for at least one time unit, i.e.,

$$W_L(0,t) = 0, \qquad 0 \leqq t < 1. \tag{3.2}$$

On the other hand, we have

$$W_L(0,t) = k_1 W_R(0, t-1), \qquad 1 \leqq t < 2, \tag{3.3}$$

since, by causality and the delay structure of the line, it will take at least two time units for any left-going incident wave to appear as an input to the first section.

Therefore we can *identify* the first reflection coefficient k_1 as

$$k_1 = \left.\frac{W_L(0,t)}{W_R(0,t)}\right|_{\text{any } t\in[1,2)}. \tag{3.4}$$

By the same argument we observe that, since it will take at least three time units for a left-going wave to be incident on the second section,

$$k_2 = \left.\frac{W_L(1,t)}{W_R(1,t)}\right|_{\text{any } t\in[2,3)}. \tag{3.5}$$

(Recall our convention, (2.20), that $W_L(1,t) = W_L(x,t)|_{x=1/2+}$.) The identification of k_1 from the scattering data is thus immediate; however we do not have at hand $W_L(1,t)$ and $W_R(1,t)$, which are needed to determine k_2. But, by using the chain-scattering (or transfer) matrix Θ_1, which is completely specified once we now know k_1, we can compute

$$\begin{bmatrix} W_R(1,t) \\ W_L(1,t) \end{bmatrix} = (1-k_1^2)^{-1/2}\begin{bmatrix} 1 & -k_1 \\ -k_1 & 1 \end{bmatrix}\begin{bmatrix} W_R(0,t-\frac{1}{2}) \\ W_L(0,t+\frac{1}{2}) \end{bmatrix} \tag{3.6}$$

and now k_2 can be computed.

In effect, once we have determined the reflection coefficient of the first section, we use its associated Θ matrix to "peel off" the effect of the first layer of the medium and put ourselves in the same position as before with a line parametrized by $\{k_2, k_3, \cdots\}$. In other words, using k_1 and the original scattering data we are able to produce the "artificial scattering data" for the line extending over $[2, \infty)$.

We can continue in this way, successively determining a reflection coefficient and peeling off the associated layer, to determine as much as we wish of the (infinite) extent of the transmission-line. For reasons to be explained in §3.2, this procedure will also be called a Schur type algorithm.

Recovery of the impedance function. Once the $\{k_i\}$ have been determined, the $\{Z_i\}$ can be found by appeal to (2.30). However, direct recovery of the impedance profile by a different layer-peeling algorithm is also possible. Causality arguments applied to the signal propagation in the gain-delay structure shown in Fig. 3(b), similar to those used in the previous section, show that we can determine Z_1 as

$$Z_1 = \left.\frac{v(1,t)}{i(1,t)}\right|_{\text{any } t\in[1,2)}. \tag{3.7}$$

Of course $v(0,t) = i(0,t)$ for $t < 1$, since $Z_0 = 1$, and using (2.28) will yield the signals at depth 1 (this initialization is in fact a trivial layer-peeling step). After determining Z_1 we can again use (2.28) to "peel off" the second medium layer and compute the scattering data $\{v(2,t), i(2,t)\}$. We have

$$\begin{bmatrix} v(2,t) \\ i(2,t) \end{bmatrix} = \Xi^{-1}(Z_1)\begin{bmatrix} \Delta & 0 \\ 0 & \Delta^{-1} \end{bmatrix}\Xi(Z_1)\begin{bmatrix} v(1,t) \\ i(1,t) \end{bmatrix}, \tag{3.8}$$

from which we obtain

$$Z_2 = \left.\frac{v(2,t)}{i(2,t)}\right|_{\text{any } t\in[2,3)} \tag{3.9}$$

and so on.

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This procedure is the discrete analogue of methods recently presented for the continuous case by Santosa and Schwetlick [23] and by Sondhi and Resnick [24] (see also [22], [25], [27]).

Remarks on numerical aspects. A count of the numerical operations in the above algorithms shows that $O(N^2)$ computations are needed for the identification process to reach a depth N on the line. The numerical properties of such recursive procedures have also been recently analyzed and the algorithms were found to be stable in the backwards sense of numerical analysis ([14], [15]).

In [14], [15] and in all our analysis so far, it was assumed that the given scattering data are exact. The effects of noisy data on the recovery of transmission-line parameters are studied in [28] (see also [29], [49]), and it is shown that the medium parameters induce a certain inherent conditioning of the inverse problem that determines the rate of error growth in the identification process. The error accumulation with depth is determined by a running product of terms of the form $(1 + |k_i|)/(1 - |k_i|)$, which shows that if the reflection coefficients are never close to 1 in absolute value, the medium identification error due to noisy data will not grow explosively (see [28]). In many practical examples we may have prior information on the range or distribution of reflection coefficients. In particular, as may often happen, if the reflection coefficients are small most of the time, the medium can be accurately recovered to a substantial depth before the error due to noisy data reaches a destructive level (see [28], [29]).

3.2. Linear fractional maps and an algorithm of I. Schur. In our previous analysis of the discrete transmission line, the impedance profile was assumed to be piecewise continuous but the input and output waveforms were functions of a continuous time parameter. In the inversion procedure we saw that only one pair of values from the scattering data is needed to determine the next reflection coefficient. This being the case we may choose the probing current waveform, $i(0, t)$, to be such that only one number per unit time will be sufficient to describe it completely, and this will then be the case for the response $v(0, t)$ too. One such choice is a piecewise constant signal; another is a sequence of impulses or, in particular, a single impulse. It is easy to realize (see e.g. Fig. 3) that the response to a piecewise constant signal is also piecewise constant, and that an impulsive forcing current will elicit a voltage response that is a weighted series of impulses, say

$$v(0,t)=\delta(t)+2\sum_{i=1}^{\infty} h_i\delta(t-i) \quad \text{when } i(0,t)=\delta(t). \tag{3.10}$$

Converting this data to right- and left-propagating wave variables, we have the following scattering data pair:

$$W_R(0,t)=\delta(t)+\sum_{i=1}^{\infty} h_i\delta(t-i) \quad \text{and} \quad W_L(0,t)=\sum_{i=1}^{\infty} h_i\delta(t-i). \tag{3.11}$$

Since the transmission-line sections operate linearly on the propagating signals, the impulse response data in (3.10), or equivalently in (3.11), will completely characterize the behavior of the semi-infinite line. (In fact, from (3.11) we can also obtain a causal impulse response that relates the reflected wave $W_L(0, t)$ to the probing wave $W_R(0, t)$. This function will play an important role later on.) In the sequel we shall therefore consider the nonuniform line as a discrete, time-invariant linear system acting on sequences of numbers, bearing in mind that we in fact operate on sequences of impulses or, alternatively, on piecewise constant waveforms.

We formally define the corresponding z-transforms of the time sequences arising in (3.10) and (3.11) as

$$i(0,z)=1, \qquad v(0,z)=1+2\sum_{i=1}^{\infty} h_j z^{-j}=1+2H(z), \tag{3.12}$$

say, and accordingly

$$W_R(0,z)=1+H(z) \quad \text{and} \quad W_L(0,z)=H(z). \tag{3.13}$$

We next introduce a function often encountered in the scattering literature, the *depth-n left reflection function*, defined as

$$R_n(z)=\frac{W_L(n,z)}{W_R(n,z)}. \tag{3.14}$$

Therefore, $R_n(z)$ is the z-transform of the left-propagating response to a probing, right-propagating impulse (i.e. the W_L/W_R transfer function), for the transmission line extending from the beginning of section n to infinity (see Fig. 4). In particular, for $n=0$, we have the wave-domain impulse response, or reflection function

$$R_0(z)=\frac{W_L(0,z)}{W_R(0,z)}=\frac{H(z)}{1+H(z)}, \qquad H(z)=\sum_{j=1}^{\infty} h_j z^{-j}. \tag{3.15}$$

Let us now try to relate the elements of the sequence of functions $\{R_n(z)\}$ to the parameters of the transmission line. Using the evolution equations (2.21) in the transform domain, or directly referring to Fig. 4, we obtain, after some easy algebra, the recursion

$$R_n(z)=\frac{zR_{n-1}(z)-k_n}{1-k_n zR_{n-1}(z)}. \tag{3.16}$$

This recursion, which is a linear fractional map, or a discrete Riccati equation, provides the evolution of the reflection function as we proceed deeper and deeper along the transmission line, provided we are given the reflection coefficients k_n. From the discussion at the beginning of this section, we know that we can determine the required reflection coefficients recursively as we propagate (3.16); merely note that k_n can be determined from past results as

$$k_n=\lim_{z\to\infty} zR_{n-1}(z). \tag{3.17}$$

For example, we have that

$$k_1=\lim_{z\to\infty} zR_0(z)=\frac{W_L(0,1)}{W_R(0,0)} \tag{3.18}$$

because the first nonzero response to an impulsive probing function $W_R(0,t)$ will be an impulse with amplitude k_1. The same is true at any depth, and therefore we have an alternative inversion algorithm, which recursively propagates (3.16) together with the identification formula (3.17). However, it can be verified that, when translated in terms of time-series coefficients, the linear fractional recursion (3.16) is simply a functional description of the steps in the layer-peeling algorithm described in §3.1. More precisely, we can "linearize" the fractional map (3.16) by determining interlocked recursions for the numerator and denominator of the function $R_n(z)$ (put $R_n(z)=U_n(z)/V_n(z)$, and using (3.16) it is easy to see that U_n and V_n obey the recursion (2.21), with $z^{-1/2}\leftarrow\Delta$).

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

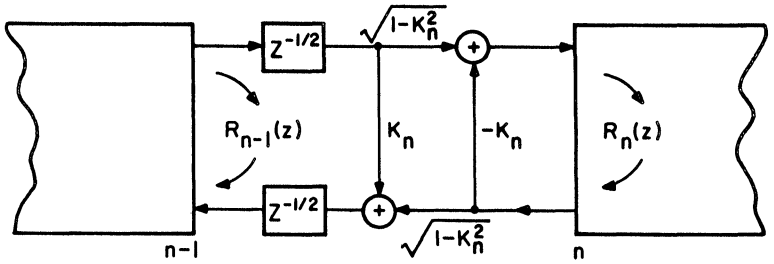

FIG. 4. *The definition and evolution of the left-reflection functions.*

The continuous counterpart of the inverse scattering algorithm (3.16), (3.17) involves the propagation of a nonlinear Riccati equation together with a boundary condition for the identification of the local reflectivity function. This method for solving continuous inverse scattering problems was first derived independently by Bruckstein, Levy and Kailath [25] and by Corones, Davison and Krueger [50]. Numerical implementations of the nonlinear recursions for continuous case "reflection kernels" $R_x(t)$ are clearly possible (see e.g. [50]). However, these recursions have a complexity of $O(N^3)$, which is, quite unnecessarily, as high as a nonefficient approach via matrix equations. This happens because, to propagate the reflection kernels, implicit deconvolutions (to provide the impulse response of the medium at depth x) have to be performed. The linearized layer-peeling methods, on the other hand, use the medium model directly to compute the *waves* propagating into the medium, and, as we have seen, these waves provide information on the local reflectivity in a more direct and efficient way.

Let us now define a power series (in z^{-1}) as

$$S_n(z) = zR_n(z) = s_0^n + s_1^n z^{-1} + \cdots \tag{3.19}$$

so that the recursion (3.16) becomes

$$S_n(z) = z\frac{S_{n-1}(z) - k_n}{1 - k_n S_{n-1}(z)}. \tag{3.20}$$

The reason for doing this is that (3.20) then turns out to be *identical* to a recursion presented by I. Schur in 1917 for checking whether a power series in z^{-1}, $S(z)$, is bounded by one outside the unit circle of the complex plane, i.e.,

$$|S(z)| < 1 \quad \text{for } |z| > 1. \tag{3.21}$$

Schur's test (see [33], [34]) is precisely that the sequence of numbers k_n implicitly defined by the recursion (3.20), and

$$\lim_{z\to\infty} S_{n-1}(z) = k_n \tag{3.22}$$

should satisfy

$$|k_n| < 1. \tag{3.23}$$

This result also shows that the left reflection function associated with any transmission-line will be bounded by one outside the unit circle of the complex plane, a fact often derived differently in the scattering literature.

The algorithm of Schur has been encountered in many recent engineering analyses, e.g., in network and digital filter synthesis, and in the derivation of fast

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

algorithms for linear estimation and fast Cholesky factorization of Toeplitz and related matrices [35]–[37], [39], [51]–[53]. Our discussion in §§2 and 3 shows that this algorithm may be regarded as describing the propagation of waves along a passive transmission line, and consequently as providing a direct inverse scattering algorithm for determining the parameters of the line given the special input and output sequences (3.13), $1 + \sum_1^\infty h_j z^{-j}$ and $\sum_1^\infty h_j z^{-j}$. We may remark that this is an important type of scattering data, being encountered in the so-called "marine seismograms," where the medium has a perfect reflection at the boundary (for example, the sea-air interface) (see Fig. 5).

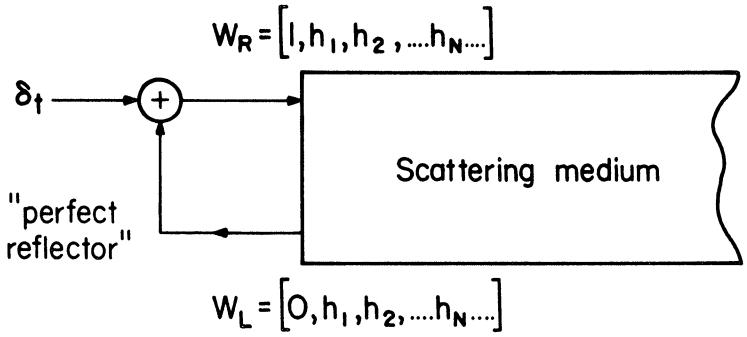

FIG. 5. *The perfect reflection experiment.*

Despite the prevalence of the special scattering data (3.13), we should stress that the inversion procedures of §3.1 work with *arbitrary* input and output sequences $W_R(0, z)$ and $W_L(0, z)$; of course, by linearity, there must exist a power series $f(z)$ (in z^{-1}) such that

$$W_R(0,z)=f(z)(1+H(z)), \qquad W_L(0,z)=f(z)H(z) \tag{3.24}$$

corresponding to exciting the medium not with a current impulse $i(0, z) = 1$ but with the function $f(z)$. In works on inverse scattering based on the classical linear equation methods (see §4), it is sometimes advocated that a "deconvolution" step be first performed to remove the effect of the excitation $f(z)$ and go back to the situation of an impulse input. Our layer-peeling formulation shows clearly that such a preliminary deconvolution is unnecessary. This is also clear from the linear fractional form (3.16) of the Schur recursion, since $R_0(z)$ and $S_0(z)$ are the same no matter what the non-zero function $f(z)$ is. We remark that, in the application to matrix factorization results, the above remarks provide a nice interpretation of certain congruence relationships between Toeplitz matrices and matrices of displacement inertia $\{1, 1\}$ (see [47], [48]).

Impedance-domain recursions. We note that we may associate with the voltage-current evolution equations (2.28) a different linear fractional transformation. Indeed, defining the *depth-n impedance function*

$$A_n(z)=\frac{v(n,z)}{i(n,z)}, \tag{3.25}$$

we obtain for its evolution the recursion

$$A_{n+1}(z)=\frac{(1+z)A_n(z)+(1-z)Z_n}{(1-z)Z_n^{-1}A_n(z)+(1+z)}. \tag{3.26}$$

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This functional recursion, together with the local impedance identification formula

$$Z_n = \lim_{z\to\infty} A_n(z), \tag{3.27}$$

provides an alternative inversion process, similar to the Schur recursions. Clearly we have the initialization $A_0(z) = 1 + 2H(z)$, and the above algorithm implicitly tests whether this function has *positive* real part outside the unit circle, since this is the property equivalent to the boundedness of $R_0(z)$ (see e.g. [34]). The function is positive real provided the Z_n sequence implicitly defined via (3.26) and (3.27) is strictly positive and bounded. We note that this result, a corollary of the boundedness test of Schur, is apparently new.

3.3. Extensions to other types of media. Up to this point we assumed that the underlying scattering medium is a piecewise constant and lossless transmission-line model. However, it is not hard to see that the layer-peeling algorithm can be used whenever each section of the line is uniquely specified by the local left reflection coefficient, without necessarily being lossless. That is, the scattering matrix of each section could have the general form

$$\Sigma_G(k) = \begin{bmatrix} \sigma_{11}(k) & \sigma_{12}(k) \\ k & \sigma_{22}(k) \end{bmatrix} \tag{3.28}$$

with σ_{ij} arbitrary and σ_{22} nonzero. The invertibility of σ_{22} is required for the existence of the chain-scattering or transfer matrix representation $\Theta_G(k)$ of the above scattering matrix. This matrix is given by the following (exchange) rule which can be obtained by some simple algebra (see e.g. [46] and Appendix A)

$$\Theta_G(k) = \begin{bmatrix} \sigma_{11}(k) - k\sigma_{22}^{-1}(k)\sigma_{12}(k) & \sigma_{22}^{-1}(k)\sigma_{12}(k) \\ -k\sigma_{22}^{-1}(k) & \sigma_{22}^{-1}(k) \end{bmatrix}. \tag{3.29}$$

In this context, we do not generally have lossless propagation; neither can we expect a simple transmission-line model to correspond to the wave propagation equations. We are merely solving an inverse scattering problem for a more complicated two component difference equation, through the solution of an initial value problem that involves propagation of the model equations together with a formula yielding the next required parameters from the already computed quantities. In this context, it is not known in general how to derive Gelfand–Levitan or Krein or Marchenko type matrix equations-based approaches (for reasons that will be evident in §4).

A particular case (3.28) of interest is the simple propagation model described by

$$\Theta_k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}.$$

This yields an input-output map, specified by the impulse response $R_0 = \sum_1^\infty k_i z^{-1}$, and the layer-peeling inverse scattering algorithm then reduces to a simple recursive deconvolution process. Another propagation model that is quite interesting arises in the solution of the partial realization problem. The minimal partial realization problem is to determine a lowest order linear system whose impulse response matches a given sequence up to a certain lag, or depth. Clearly we could realize any sequence by an almost arbitrary propagation model; however, the *lowest order* requirement forces the choice of a propagation model with a minimal number of delay sections. In [30], inverse scattering principles, associated with a slightly more general layered model, were applied to analyze the partial realization problem. The layer-peeling idea then directly led to a relatively recent partial realization procedure, the generalized

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Lanczos algorithm; the inverse scattering framework also provided several alternative solutions. Those subsequently found an application in devising efficient architectures for decoding Reed–Solomon and BCH error-correcting codes (see [31]).

Analyzing layer-peeling inverse scattering algorithms, it becomes clear that the crucial property allowing a recursive identification of the medium is the possibility to determine the nearest scattering layer from a causal input–output pair recorded at its left boundary. In gain-delay type media such as the ones encountered above, we see that the initial portion of the response can be attributed solely to the action of the nearest layer and this is sufficient for its identification.

To make an even stronger case for the easily derived and understood layer-peeling methods, we may note that it will work even on a nonlinear gain-delay type scattering model with a special structure. Indeed, suppose we are given a model for the local behavior that has alternating delay sections with nonlinear interactions, described by the scattering operator

$$\begin{bmatrix} W_R(n+1,t) \\ W_L(n,t) \end{bmatrix} = \begin{bmatrix} F^t_{k_n}\{W_R(n,t-\frac{1}{2})\} + F^\rho_{k_n}\{W_L(n+1,t)\} \\ F^r_{k_n}\{W_R(n,t-1)\} + F^\tau_{k_n}\{W_L(n+1,t-\frac{1}{2})\} \end{bmatrix} \tag{3.30}$$

where $F^t_k\{\cdot\}$, $F^\tau_k\{\cdot\}$, $F^r_k\{\cdot\}$ and $F^\rho_k\{\cdot\}$ are static nonlinear functions, parametrized by k and assumed to have the following properties.

1) They all pass through the point (0, 0) for all k, i.e., $F_k(0) = 0$;
2) $F^\tau_k\{\cdot\}$ is an invertible transformation for all k;
3) Given v and the value of $F^r_k\{v\}$, k is uniquely determined.

With these assumptions we have the following straightforward inversion algorithm:

(a) From $F^r_{k_n}\{W_R(n, n/2)\} = W_L(n, n/2 + 1)$ determine k_n.
(b) Compute the waves at $n + 1$ using first

$$W_L(n+1,t) = F^{\tau^{-1}}_{k_n}\{W_L(n,t+\tfrac{1}{2}) - F^r_{k_n}\{W_R(n,t-\tfrac{1}{2})\}\} \tag{3.31a}$$

then

$$W_R(n+1,t) = F^t_{k_n}\{W_R(n,t-\tfrac{1}{2})\} + F^\rho_{k_n}\{W_L(n+1,t)\}. \tag{3.31b}$$

Note that (3.31a, b) effectively yield the nonlinear transfer operator corresponding to the scattering layer described by (3.30).

(c) Return to step (a) with the forward propagated scattering data $W_R(n + 1, t)$ and $W_L(n + 1, t)$.

The above algorithm is guaranteed to proceed successfully provided the scattering data was genuine, i.e., in fact produced by the model (3.30). The algorithm above may be viewed as an example of nonlinear inverse scattering via downward continuation. We also note that in the nonlinear case we do not have an equivalent fractional transformation-based method, recursively propagating for the impulse responses of the media considered over the intervals $[n, \infty)$. In fact the nonlinear system described above *is* completely determined by its impulse because, as we saw above, from it the medium parameters can be recovered. However, in order to obtain the response of the medium to an arbitrary input, one has no alternative but to propagate this input through the scattering structure.

At a first glance, the above generalization might seem rather unmotivated. However some recent papers (see e.g. [32]) addressed the issue of numerically solving nonlinear Volterra equations of the convolution type, which appear in modeling biological and neurophysiological processes. The discretized version of such integral

equations can be written in our notation as follows:

$$W_L(0,t)=\sum_0^t F\{k_{t-\xi}, W_R(\xi)\} \quad \text{or} \quad =\sum_0^t F\{k_\xi, W_R(t-\xi)\}. \tag{3.32}$$

It is not hard to realize that the above equation describes the input/output relation of a nonlinear scattering medium with the following scattering description

$$\begin{bmatrix} W_R(n+1,t) \\ W_L(n,t) \end{bmatrix}=\begin{bmatrix} W_R(n,t-1) \\ F\{k_n, W_R(n,t-1)\}+W_L(n+1,t) \end{bmatrix}, \tag{3.33}$$

which is a particular case of (3.30). Therefore the solution of (3.32), i.e., the determination of the function k_n for all n, can proceed, under some mild conditions on the nonlinear function involved, via a fast layer-peeling algorithm. In the analysis done by Hairer et al. [32], the conclusion is that some even faster ($O(N\log^2 N)$) algorithms can be obtained to solve nonlinear convolution equations of the Volterra type. Very fast algorithms already exist for the linear deconvolution problem ($O(N\log N)$, via the fast Fourier transform technique). More importantly, several recently obtained results for fast factorizations of structured matrices [53]–[55] also indicate that such $O(N\log^2 N)$ algorithms can be devised for the inverse scattering problem too, via a so-called *doubling* technique. Although we shall not discuss this issue in detail, we briefly note that the idea is the following. Suppose we only use N lags of the scattering data to recover the medium up to depth N. The identified portion of the medium is a linear system relating the original scattering data to the waves at depth N in the medium, via a matrix transfer function (corresponding to the cascade of the N determined sections of the scattering medium (see §4)). Thus to compute the waves at depth N we have to perform convolutions of the original scattering data with the impulse response matrix of the already determined layers. The doubling step is to use a portion of the medium of depth N to propagate for only $2N$ lags of the signals at depth N, and to do the convolutions involved via the fast Fourier transform. The first N lags of the waves at depth N will, of course, be zero. By causality, however, the next N lags will be a causal pair of scattering data for the portion of the scattering medium starting at depth N. Therefore one can now identify the medium for N additional steps in depth, use the $2N$ portion of the medium to propagate for $4N$ lags of the signals at depth $2N$ and so on. A complexity count for this process quickly yields that the number of operations required is $O(N\log^2 N)$.

We also mention that further generalizations of layer peeling procedures to gain-delay type media with n-dimensional vectors as outputs and reflected waves are possible and, implicitly, have found applications in designing digital filters well matched to VLSI implementations and to the development of fast, structured estimation algorithms for nonstationary processes [39], [51]–[53].

4. Linear equation methods for inverse scattering. The classical approaches to the solution of inverse scattering problems are via sets of matrix equations with increasing dimensionality. In this section we show that all the classical methods can be derived by a straightforward analysis of the matrix transfer functions corresponding to cascades of gain-delay sections. We shall first derive, using causality and symmetry arguments, an apparently new, general linear equation, specializations of which for different types of scattering data yield the various previously used equations.

Suppose we have the matrix transfer function corresponding to the cascade of the first n sections of the transmission-line. Therefore we have, using the z-transform

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

notation,

$$\begin{bmatrix} W_R(n,z) \\ W_L(n,z) \end{bmatrix} = \begin{bmatrix} m_{11}(n,z) & m_{12}(n,z) \\ m_{21}(n,z) & m_{22}(n,z) \end{bmatrix} \begin{bmatrix} W_R(0,z) \\ W_L(0,z) \end{bmatrix} \tag{4.1}$$

where recall that, for simplicity, we denote by $W_{R/L}(n, z)$ the sequence of values of the wave variables at depth $x = (n/2)+$. The $\{m_{ij}\}$ are sometimes known as the "influence" or Green functions associated with the propagation equations (2.14). It readily follows from the structure of the transmission line that

$$\begin{bmatrix} m_{11}(n,z) & m_{12}(n,z) \\ m_{21}(n,z) & m_{22}(n,z) \end{bmatrix} = \Theta(k_n)\Delta_S(z)\cdots\Theta(k_1)\Delta_S(z) = \prod_{i=n}^{1} \{\Theta(k_i)\Delta_S(z)\} \tag{4.2}$$

where we have used the following z-transform representation for the delay operator:

$$\Delta_S(z) = \begin{bmatrix} z^{-1/2} & 0 \\ 0 & z^{1/2} \end{bmatrix}. \tag{4.3}$$

There is complete symmetry in the forward transmission-line picture of Fig. 2(a), in that we can replace Δ by Δ^{-1} and W_R by W_L without affecting the relationships between the propagating signals. We therefore have the following useful identities:

$$m_{22}(n,z) = m_{11}(n,z^{-1}) \quad \text{and} \quad m_{12}(n,z) = m_{21}(n,z^{-1}). \tag{4.4}$$

At this point a cosmetic step of redefining the unit delay is helpful in simplifying the subsequent arguments. Setting $z^{-1} \leftarrow z^{-1/2}$ in (4.3) redefines the time scale to $t \leftarrow 2t$. Therefore we shall regard, conceptually, the time functions as sequences of numbers padded with zeros at alternate time lags. Because of the structure of the "relative shift" matrix

$$\Delta_S(z) = \begin{bmatrix} z^{-1} & 0 \\ 0 & z \end{bmatrix}, \tag{4.5}$$

which contains both time-delay and time-advance operators, the time support of the functions/sequences $\{m_{ij}(n, z)\}$ extends from $t = -n$ to time $t = n$ (in terms of the rescaled time), with exactly $n + 1$ nonzero lags. Figure 6 shows the support of the signals of interest in the (x, t) plane. Also note that by causality, at depth $x = n/2$ we have (see Fig. 2(b))

$$W_R(n,t) = 0 \quad \text{for } t < n, \qquad W_L(n,t) = 0 \quad \text{for } t < n+1 \tag{4.6}$$

while for $t = n$ we have

$$W_R(n,n) = \prod_{i=n}^{1} (1-k_i^2)^{1/2} W_R(0,0). \tag{4.7}$$

We can rewrite the basic relation (4.1) in the time domain as

$$\begin{aligned} W_R(n,t) &= W_R(0,t) * m_{11}(x,t) + W_L(0,t) * m_{12}(x,t), \\ W_L(n,t) &= W_R(0,t) * m_{21}(x,t) + W_L(0,t) * m_{22}(x,t) \end{aligned} \tag{4.8}$$

where $*$ stands for (sequence) convolution. Now define the function

$$\Psi(n,t) = m_{11}(n,t) + m_{21}(n,t), \tag{4.9}$$

which has a time span of $[-n, n]$. Because of (4.4), it turns out that

$$\Psi(n,-t) = m_{22}(n,t) + m_{12}(n,t). \tag{4.10}$$

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

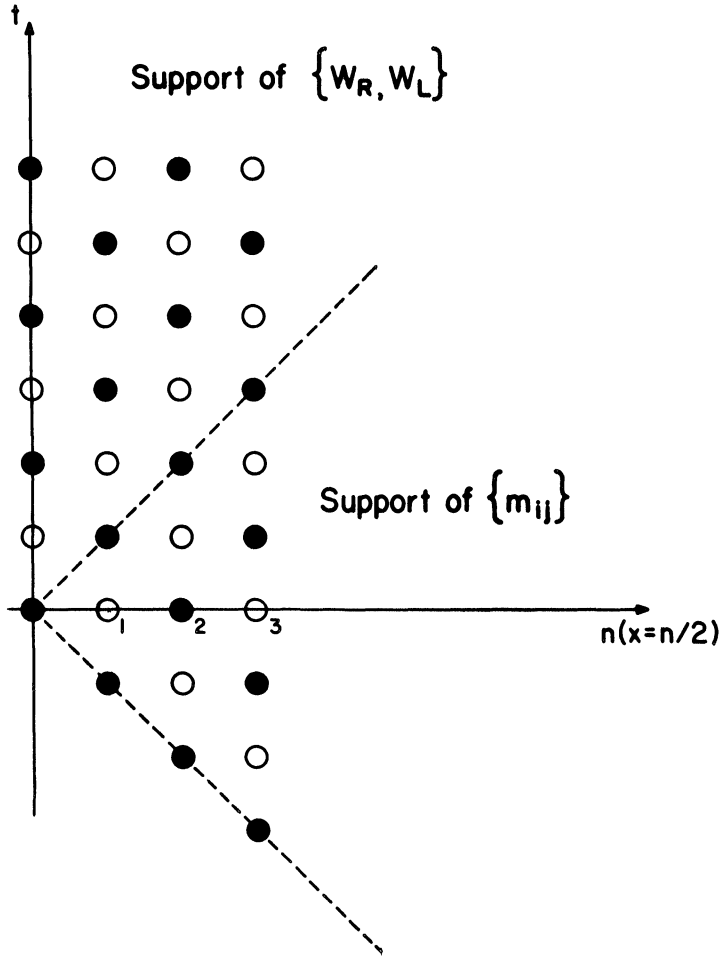

FIG. 6. *Support of influence functions and scattering data in the space-time plane.*

Therefore, adding the two equations of (4.8), we have

$$W_R(0,t) * \Psi(n,t) + W_L(0,t) * \Psi(n,-t) = W_R(n,t) + W_L(n,t). \tag{4.11}$$

Writing out the convolutions explicitly for $t \leqq n$, and using the causality relations (4.6), we obtain

$$\sum_{-n}^{n} W_R(0,t-i)\Psi(n,i) + \sum_{-n}^{n} W_L(0,t-i)\Psi(n,-i) = \begin{cases} 0 & \text{for } t<n, \\ W_R(n,n) & \text{for } t=n. \end{cases} \tag{4.12}$$

In (4.12) the signals $\{W_R(0,t), W_L(0,t)\}$ are causal funtions (i.e., their support is $t \geqq 0$). Therefore we can rewrite the above result as the following set of linear equations:

$$\{\mathbf{L}_T[W_R^n] + \mathbf{L}_T[W_L^n]\bar{I}\}\Psi^n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ W_R(n,n) \end{bmatrix} \tag{4.13}$$

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where

1) Ψ_n is a column vector of dimension $n+1$ stacking all the nonzero elements of the sequence $\Psi(n, t)$ from the span $t \in [-n, n]$ in increasing order of the time index;
2) $\mathbf{L}_T[X]$ is a lower triangular Toeplitz matrix with first column X;
3) $W^n_{R/L} = [W_{R/L}(0, 0)\ W_{R/L}(0, 1) \cdots W_{R/L}(0, n)]^T$;
4) $\bar{I}$ is a matrix with ones on the antidiagonal and acts as a time reversal operator.

The above general linear equation was not derived earlier in the inverse scattering literature. Usually, equations corresponding to special choices of scattering data sequences $\{W_R(0, t), W_L(0, t)\}$ (or corresponding special choices of $\{i(0, t), v(0, t)\}$) are obtained by various, specifically tailored approaches that successfully conceal the general causality structure leading to the result above. Let us now proceed to derive from (4.13) the classical linear equations that appear in the inverse scattering literature.

4.1. The Gelfand–Levitan, Marchenko and Krein equations. Suppose that the scattering data corresponds to the current impulse response, i.e. that we are given

$$i(0,t)=\delta(t), \qquad v(0,t)=1+2\sum_1^\infty h_i\delta(t-i) \tag{4.14}$$

or correspondingly

$$W_R(0,t)=\delta(t)+\sum_1^\infty h_i\delta(t-i), \qquad W_L(0,t)=\sum_1^\infty h_i\delta(t-i). \tag{4.15}$$

Then (4.13) reduces to the form

$$\{\mathbf{L}_T[E+H^n]+\mathbf{L}_T[H^n]\bar{I}\}\Psi^n=\begin{bmatrix}0\\0\\\vdots\\\prod_n^1(1-k_i^2)^{1/2}\end{bmatrix}, \tag{4.16}$$

a linear system of equations with a Toeplitz + Hankel coefficient matrix. This is a discretized version of the Gelfand–Levitan equation. Note that we defined H^n to be $[0, h_1, h_2, \cdots, h_n]^T$ and $E = [1, 0, \cdots, 0]^T$.

Multiplying both sides of (4.16) by $\bar{I}$, adding the direct and time-reversed equations and then using some simple algebraic tricks, such as the simple observation that $(I+\bar{I})(I+\bar{I}) = 2(I+\bar{I})$, results in a completely symmetric equation for the vector $\Psi_n + \bar{I}\Psi^n$ which corresponds to $\Psi(n, t) + \Psi(n, -t)$. This equation is

$$\left\{I+\frac{1}{2}(I+\bar{I})\mathbf{L}_T(H^n)(I+\bar{I})\right\}[\Psi^n+\bar{I}\Psi^n]=\begin{bmatrix}\prod_n^1(1-k_i^2)^{1/2}\\0\\\vdots\\0\\\prod_n^1(1-k_i^2)^{1/2}\end{bmatrix}, \tag{4.17}$$

and it should be considered as the "true" discretized Gelfand–Levitan system, since the continuous counterpart also has a completely symmetric kernel (see e.g. [25]).

Assume next that the scattering data is the left reflection function $R_0(z)$, or equivalently that we are given

$$W_R(0,t)=\delta(t) \quad \text{and} \quad W_L(0,t)=\sum_{i=1}^\infty r_i\delta(t-i). \tag{4.18}$$

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Substituting this into the general equation (4.13), leads to

$$\{I+\mathbf{L}_T[R^n]\bar{I}\}\Psi^n=\begin{bmatrix}0\\ \vdots\\ \prod_n^1(1-k_i^2)^{1/2}\end{bmatrix} \tag{4.19}$$

with $R^n=[0, r_1, r_2 \cdots r_n]^T$, which can be recognized as a discrete Marchenko equation (see e.g. [17]).

The Gelfand–Levitan and Marchenko equations have coefficient matrices that are respectively Toeplitz + Hankel and Hankel matrices. There is a somewhat lesser known formulation in the inverse scattering literature due to Krein (see [3], [5], [25], [38]), which has a purely Toeplitz cofficient matrix. To derive this equation we use the special scattering data sequences that led to the Gelfand–Levitan equation

$$W_R(0,t)=\delta(t)+\sum h_i\delta(t-i), \qquad W_L(0,t)=\sum h_i\delta(t-i) \tag{4.20}$$

and rewrite (4.8) as

$$\begin{aligned} W_R(n,t)&=m_{11}(n,t)+(\textstyle\sum h_i\delta(t-i))*[m_{11}(n,t)+m_{12}(n,t)],\\ W_L(n,t)&=m_{21}(n,t)+(\textstyle\sum h_i\delta(t-i))*[m_{11}(n,t)+m_{12}(n,t)]. \end{aligned} \tag{4.21}$$

If we now define

$$\Phi(n,t)=m_{11}(n,t)+m_{12}(n,t)=m_{22}(n,-t)+m_{21}(n,-t), \tag{4.22}$$

then write the second equation above for reversed time and add these equations yielding $W_R(n,t)$ and $W_L(n,t)$ over $t\leqq n$, we obtain the Krein equation

$$\{I+\mathbf{L}_T[H^n]+\bar{I}\mathbf{L}_T[H^n]\bar{I}\}\Phi^n=\begin{bmatrix}0\\ \vdots\\ \prod_n^1(1-k_i^2)^{1/2}\end{bmatrix}. \tag{4.23}$$

This equation has a symmetric Toeplitz matrix of coefficients, defined by the sequence $[1, h_1, h_2, \cdots, h_n]$, that will be denoted by $\mathbf{S}_T[E+H_n]$. Recall that in deriving (4.23) we again used the causality relations (4.6).

4.2. Fast algorithms for matrix equation based approaches. To obtain the inverse scattering methods associated with the above derived sets of linear equations, in forms commonly encountered in the literature, e.g., [6], [7], [13], [16], [17], we have to convert the left-hand side of the above equations to $[0\ 0 \cdots 1]^T$, which is independent of the reflection coefficient to be recovered. This is done by multiplying both sides by $W_R(n,n)=\prod_1^n(1+k_i^2)^{-1/2}$. In this case we have to solve the matrix equations for the vectors $\Phi_n\prod_1^n(1-k_i^2)^{-1/2}$ and $\Psi_n\prod_1^n(1-k_i^2)^{-1/2}$. To recover the parameters of the medium from the solutions of these equations we may then invoke two identities,

$$\prod_n^1(1-k_i^2)^{-1/2}\sum_{-n}^n\Psi_n=\prod_n^1\frac{1}{1+k_i}=\prod_n^1(1-k_i^2)^{-1/2}\sum_{-n}^n\Phi_n, \tag{4.24}$$

for which we give a very simple proof in Appendix B. From these results it becomes obvious that the ratio of the sums of entries for consecutive solutions readily provides the sequence of reflection coefficients.

We also note that a multiplication of both sides of the discrete Krein equation by $\prod_n^1(1-k_i^2)^{1/2}$ would yield another particular form to its solution. It would have unit leading coefficient ($t=n$) and k_n for the value at $t=-n$ (see Appendix B). This

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

form of the Krein linear system of equations is well known in linear prediction theory as the set of "normal equations" for the predictor coefficients. The solution of such equations via the fast, $O(N^2)$, Levinson–Durbin recursive method opened up the field of fast lattice algorithms in estimation theory, (see e.g. [47], [51]–[55]). Among other things, research in this field provided fast algorithms for classes of near-Toeplitz and near-Hankel matrices, which themselves have interesting connections with transmission-lines (see e.g. [51], [52]).

To derive *fast algorithms* that recursively solve the above-derived linear equations for increasing values of n, we shall use the fact that the solution vectors are related to the transfer representations of the medium over the intervals $[0, n/2]$. In particular we have to determine

$$\Psi(n,t)=m_{11}(n,t)+m_{21}(n,t) \quad \text{and} \quad \Phi(n,t)=m_{11}(n,t)+m_{12}(n,t) \tag{4.25}$$

recursively for increasing values of n.

Recalling (4.2), we have for the entries of the transmission representation (which determine the vector solutions of the matrix equations) the recursions

$$\begin{bmatrix} m_{11}(n,z) \\ m_{21}(n,z) \end{bmatrix} = \Theta(k_n)\Delta_S(z) \begin{bmatrix} m_{11}(n-1,z) \\ m_{21}(n-1,z) \end{bmatrix} \tag{4.26}$$

with initial conditions $m_{11}(0, z) = 1$ and $m_{21}(0, z) = 0$. Suppose now that we have already computed the (influence or Green) functions $m_{11}(\cdot, t)$ and $m_{21}(\cdot, t)$ to depth $x = n/2$. These functions are not causal but have time support $[-n, n]$ at depth $x = n/2$, and therefore we have computed up to this stage about $n(n-1)/2$ numbers. Together with the scattering data, the Green functions at n provide, via a convolution operation, the causally propagated waves $W_R(n, t)$ and $W_L(n, t)$. To proceed with the recursions above, we need the value of k_{n+1} which (recalling the results of §3.1) is given by

$$k_{n+1}=\frac{W_L(n,n+1)}{W_R(n,n)}=\frac{[W_R(0,t)*m_{21}(n,-t)+W_L(0,t)*m_{11}(n,-t)]_{t=n+1}}{\prod_n^1(1-k_i^2)^{1/2}}. \tag{4.27}$$

In order to recursively compute the vectors involved in the equations of inverse scattering, we therefore need to compute at each depth two inner products (convolutions). The resulting algorithms, when based on (4.26), require about $4n$ multiplications to go from depth n to depth $n+1$, and therefore the recursive solution of the matrix equations of inverse scattering will take $O(N^2)$ time.

The above approach is general enough to work for all sets of scattering data. In the Toeplitz (or normal equations) case, corresponding to the data $H(z)$, the resulting algorithm is the well-known Levinson–Durbin recursion. Indeed, note that when $W_R(0, t) = W_L(0, t)$ for $t \geqq 1$, we have by (4.27) that only one inner product has to be computed to determine the next reflection coefficient and proceed with the algorithm. This is indeed what the Levinson–Durbin algorithm does.

In the inverse scattering literature Berryman and Greene obtained a fast algorithm providing the solutions of the discrete Marchenko equation [17]. Their algorithm, modeled on the Levinson–Durbin procedure, was derived by exploiting the Hankel structure of the coefficient matrix. Both these algorithms are now seen to be particular cases of our general method compactly described by (4.26) and (4.27).

We note again that the above presented general structure was not recognized in previous work on inverse scattering algorithms. As a result the fast algorithms that were derived for the various matrix-equations-based identification methods were only

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

considered as clever ways to exploit the structure of the matrices involved, rather than a direct consequence of the scattering medium structure.

It is also interesting to point out that, while the Schur-type algorithms can be viewed as *layer identification and peeling* processes, the linear equations based methods should be regarded as methods of recursively building up the medium representations for increasing depth by *adjoining successively identified layers* (Fig. 7). In this sense the latter methods are duals of the Schur type procedures. We note that if we are only interested in an inverse scattering method, the Schur recursions are preferable, since they work directly on the scattering data in a simple way and provide the reflection coefficients with one division, with no need to compute inner products. This latter property is a definite advantage when parallel processing of the data is possible. In parallel implementations, computing the sums of the inner products (as, e.g., in (4.27)) is a processing-time bottleneck and even with N processors available, the inversion via the fast matrix-equations algorithms takes $O(N \log N)$ time. The layer-peeling algorithms do not have inner products and can thus yield an $O(N)$ time speedup with an array N processors (see e.g. [31], [61]). As discussed in [28], the layer peeling algorithms are also better suited for intelligently incorporating prior information to improve the performance of the medium identification process with noisy data.

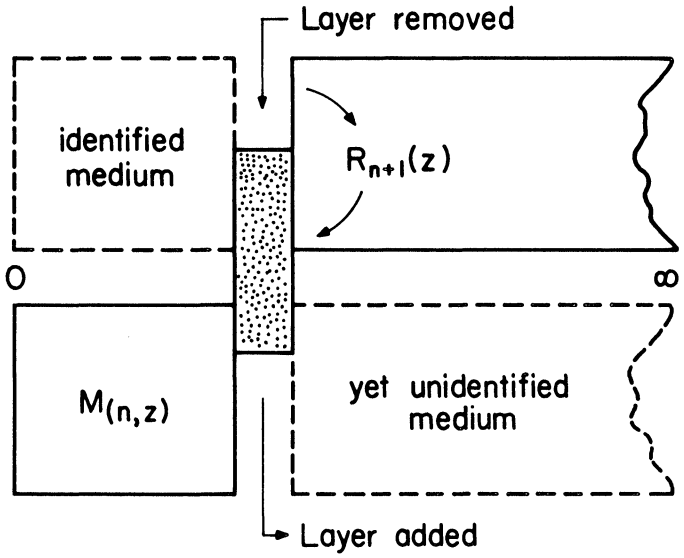

FIG. 7. *Layer-peeling vs. layer-adjoining for fast inversion algorithms.*

We also note that the doubling algorithm for inverse scattering discussed in §3.3 computes the Green functions $m_{ij}(n, t)$ for $n = 1, 2, 4, 8 \cdots$ and so on, and in fact may be viewed as an accelerated layer adjoining procedure. This algorithm, we recall, has a computational complexity of $O(N \log^2 N)$.

It should be clear from the above discussion of the special Gelfand–Levitan, Marchenko and Krein equations that a host of special equations can be obtained by various choices of $\{W_R(0, t), W_L(0, t)\}$. Indeed, several papers in the literature address the case of step response or even ramp response data. We shall not pursue this here, but, because they are widely known, we further discuss another type of inversion procedure, based on a particular method of solving integral equations that is also due

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

to Krein (see [62]). Such methods were first presented for the continuous case by Gopinath and Sondhi, and later by Burridge [20], [45]. Recently, Caflisch worked out the details of the Gopinath–Sondhi inversion procedure for the discrete case [63]. Our derivations, based on an adjoint lemma, seem to be much simpler and better suited to establishing the connections to the classical inversion methods discussed above.

4.3. Gopinath–Sondhi type inversion procedures. Suppose that, instead of solving the Krein equations, we obtain the vectors Ω_n that are solutions of

$$\mathbf{S}_T[E+H_n]\Omega_n=[1 \quad 1 \quad \cdots \quad 1]^T. \tag{4.28}$$

Such solutions have a special meaning in the theory of transmission-lines (as forcing functions that yield a propagating step of voltage on the line, see [40]) and also in the theory of nonuniform vibrating strings. This led Krein to consider a special method of solving Volterra type integral equations, via nested solutions having the unit function as the right-hand side (see [63]). The derivations that follow consider the discrete version of the Volterra-type integral equations first obtained by Krein and later used by Gopinath and Sondhi.

A question that arises naturally is whether we can relate, in a simple way, the solutions of (4.28) to those of the "original" Krein equation (4.23). The answer is obtained via the following simple adjoint lemma.

LEMMA. *If* $\mathbf{R}$ *is a symmetric nonsingular matrix, then we have that*

$$\mathbf{R}X_i=Y_i \quad \textit{for } i=1,2 \textit{ implies} \quad \langle X_1, Y_2\rangle=\langle X_2, Y_1\rangle \tag{4.29}$$

where $\langle\cdot,\cdot\rangle$ *denotes the usual inner product. The proof is via the elementary calculation below*

$$\langle X_1, Y_2\rangle=\langle X_1, \mathbf{R}_N X_2\rangle=\langle \mathbf{R}_N X_1, X_2\rangle=\langle Y_1, X_2\rangle=\langle X_2, Y_1\rangle.$$

Now it is easy to obtain a relation between the solutions of symmetric equations with right-hand sides $[0\ 0\ \cdots\ 1]^T$ and $[1\ 1\ \cdots\ 1]^T$. Indeed, by (4.29)

$$\langle \Omega_n, [0 \quad \cdots \quad 1]\rangle=\prod_n^1 (1-k_i^2)^{-1/2}\langle \Phi_n, [1 \quad \cdots \quad 1]\rangle,$$

which explicitly reads as

$$\Omega_n(n)=\left\{\prod_n^1 (1-k_i^2)^{-1/2}\right\}\sum_{-n}^{n}\Phi(n,i)=\prod_n^1 (1+k_i)^{-1}. \tag{4.30}$$

This expression could already be used to successively obtain the parameters $\{k_i\}$; however, in his paper presenting the Gopinath–Sondhi inverse scattering algorithm for the discrete case, Caflisch does not derive or use (4.30) but, paralleling the continuous case derivations, proceeds to define and use the so-called "central mass sequence" of Krein (see e.g. [42], [45], [63]). This is a sequence of positive numbers ρ_n, associated with the solutions Ω_n and defined as

$$\frac{1}{\rho_n}=\sum_{i=0}^{n}\Omega_n(i)=\langle \Omega_n, [1 \quad \cdots \quad 1]\rangle. \tag{4.31}$$

It is then proved, through energy considerations, that, in complete analogy with the continuous case result we have (see [63] and also [64]) that

$$\frac{1}{\rho_n}=\sum_0^n \frac{1}{Z_n} \tag{4.32}$$

from which one can recover the impedance profile $\{Z_i\}$.

We stress however that there is absolutely no need to introduce the central mass sequence in order to solve the inverse problem, since $\Omega_n(n)$ already contains all the necessary information for inversion. In fact we could derive other inversion methods based on solving nested sets of Toeplitz systems of equations with almost arbitrarily chosen left-hand side vectors (see [64], [65]).

Since the adjoint lemma (4.29) holds for any symmetric operator, we can readily derive a counterpart of the Gopinath–Sondhi method that is based on the discrete Marchenko equation. In the continuous case this was already done in a different way by Burridge [20], but for the discrete case we have not yet encountered the corresponding result. Suppose that we have the solutions Γ_n for the equations

$$\{I+\mathbf{L}_T[0, r_1, r_2, \cdots, r_n]\bar{I}\}\Gamma_n=[1 \quad 1 \quad \cdots \quad 1]^T. \tag{4.33}$$

Invoking the adjoint lemma (the matrix $\{I+\mathbf{L}_T[0, r_1, r_2, \cdots, r_n]\bar{I}\}$ is symmetric), we obtain

$$\Gamma_n(n)=\prod_n^1 (1-k_i^2)^{-1/2}\sum_{-n}^{n}\Psi(n,i)=\prod_n^1 (1+k_i)^{-1}. \tag{4.34}$$

This immediately provides an apparently new discrete inverse scattering algorithm. To go even further, we can find yet another inversion method by starting from the symmetrized Gelfand–Levitan equation. There we would obtain that the sum of the first and last elements of the solution corresponding to the $[1\ 1\ \cdots\ 1]^T$ right-hand side is a simple function of the reflection coefficients, which can therefore be recursively determined from consecutive solutions of increasing order. We shall not give the details since the process is clearly exemplified by the previous derivations.

A further discussion of the connections between the parameters appearing in the above calculations and the central mass sequence as well as the relation of inverse scattering to matrix factorization can be found in [65]. The reader can see again that diverse variations of the above procedures and identities are available. We shall not pursue any more here. However, we mention that a common link between various approaches to inverse scattering problems may be found in the fact that an inversion algorithm implicitly produces triangular factorizations for inverses of the coefficient matrices involved. The uniqueness of triangular factorization is at the heart of many of the special identities and relations between various parametrizations used in the theory of inverse scattering (see, e.g., [57], [64], [65]).

5. Concluding remarks. We have presented a unified and comprehensive view of the theory dealing with scattering of time sequences propagating in space through layered media modeled by transmission-line equations. Causality principles proved to be the crucial factor in solving the inverse scattering problem. Causality of the scattering representation of the medium layers leads to the possibility of deriving recursive layer-peeling algorithms for inversion, whereas the causal propagation through a cascade of symmetric layers is recognized as the reason for the existence of approaches based on linear equations.

The novel results that were discussed in this paper were the applicability of layer-peeling inverse scattering algorithms to certain classes of nonlossless media, a unified and simple derivation of all classical linear-equation based methods including a new general linear equation for inversion with arbitrary scattering data, and (by exploiting an adjoint lemma) a new discrete version of a Gopinath–Sondhi type inversion method. It was further stressed in this paper that the recursive layer-peeling methods of inverse scattering are efficient inversion algorithms, and that those are obtained by directly exploiting the structure of the scattering medium. We also showed that the

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

derivation of fast algorithms in connection with classical matrix-equations-based methods is not so much a result of the structure of the coefficient matrices involved as it is the direct consequence of the scattering medium structure. The derivations of the classical linear-equations-based inversion methods requires the exploitation of symmetry properties that are implicit in lossless media; and it should be clear that such equations can be derived for any media with similar properties. Both alternatives are applicable only when the medium is symmetric and suitably parametrized by the local reflection coefficient and lossless media are such. In this sense the direct, layer-peeling approaches that we presented are more general.

The problem of one-dimensional inverse scattering has been an active research subject for more than two decades already and the long list of references that follows is therefore far from being comprehensive. Many of the results presented in this paper have been discovered and often rediscovered in the various fields in which inverse problems arise: physics and mathematics, geophysics, network theory, linear estimation, acoustics and voice synthesis. Our paper offers what we hope is a simple but comprehensive synthesis and overview of many known results, and some that appear to be new, based on the understanding that rather general causality and symmetry properties of scattering media underlie all of them.

Appendix A. The basics of signal-flow graph representations. A signal flow diagram is a simple pictorial representation of relations between time functions or sequences, via a directed graph. The edges of the graph represent either a multiplication of the function by a constant gain factor (specified on the edge) or the action of an operator acting on the time-function (specified by a box with a label). The nodes of the graph are either summing points, so that the unique outgoing flow is the sum of two time-functions entering that node, or branching points in which the unique incoming function is replicated on the outgoing edges. These definitions are summarized in Diagram A.1 in Fig. 8.

The direction of flow on any edge of a diagram may be reversed together with the inversion of the operator corresponding to that edge (assuming that the operator is indeed invertible). This procedure will clearly leave the relationship between the time functions at the ends of the edge unchanged. If an inverted edge left a summing or branching node, we will also have to invert the flow on another incoming edge at that node. At a summing node we will also have to multiply by -1 the flow on the branch that still enters the node (see Diagram A.2 in Fig. 8). This leads us to the issue of inverting the flow on a directed path. The Mason rule for inverting the flow on a directed path is: invert each edge and, at summing nodes, change the sign of the incoming flow. This rule quickly proves the result on converting scattering diagrams to transfer representations and vice versa (see Diagram A.3 in Fig. 8).

Appendix B. Some properties of the Ψ and Φ functions. For the functions defined before, $\Psi(n, t)$ and $\Phi(n, t)$, we have the following results. In the transform domain

$$\begin{aligned}[\Psi(n,z) \quad \Psi(n,z^{-1})] &= [1 \quad 1]\begin{bmatrix} m_{11}(n,z) & m_{21}(n,z^{-1}) \\ m_{21}(n,z) & m_{11}(n,z^{-1}) \end{bmatrix} \\ &= \left(\frac{1-k_n}{1+k_n}\right)^{1/2} [z^{-1} \quad z]\begin{bmatrix} m_{11}(n-1,z) & m_{21}(n-1,z^{-1}) \\ m_{21}(n-1,z) & m_{11}(n-1,z^{-1}) \end{bmatrix}\end{aligned} \tag{B.1}$$

setting $z = 1$ leads to

$$\Psi(n, z=1) = \sum_{t=-n}^{n} \Psi(n,t) = \prod_{n}^{1} \left(\frac{1-k_i}{1+k_i}\right)^{1/2} = Z_n^{-1/2}. \tag{B.2}$$

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

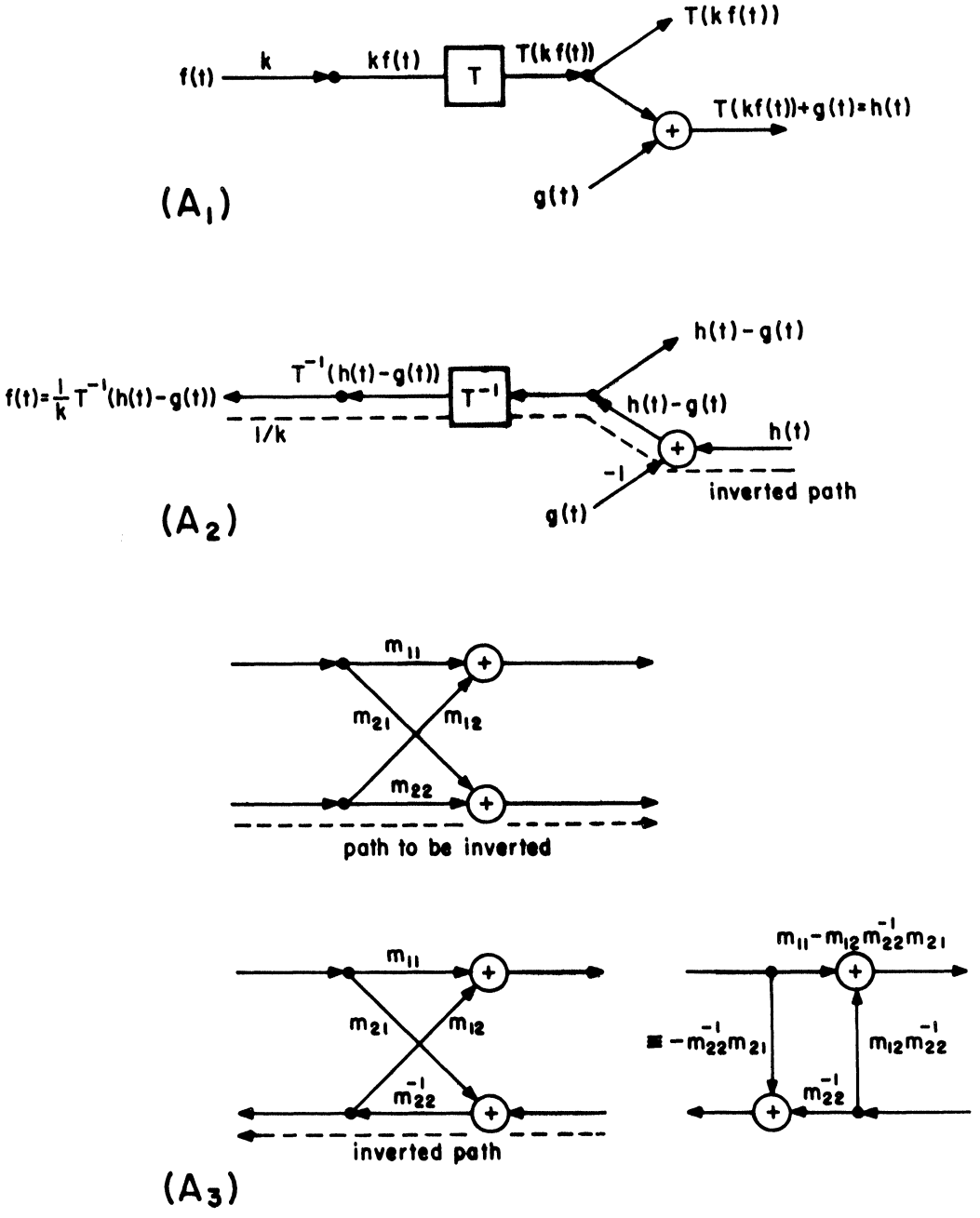

FIG. 8

For the alternate functions $\Phi(n, t)$ we have

$$\begin{bmatrix} \Phi(n,z) \\ \Phi(n,z^{-1}) \end{bmatrix} = \begin{bmatrix} m_{11}(n,z) & m_{12}(n,z) \\ m_{12}(n,z^{-1}) & m_{11}(n,z^{-1}) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{B.3}$$

This provides the recursion

$$\Phi(n,z) = (1-k_n^2)^{-1/2}[z^{-1}\Phi(n-1,z) - k_n z\Phi(n-1,z^{-1})] \tag{B.4}$$

which, after some easy algebraic manipulations, yields the following results for $\Phi(n, t)$:

$$\Phi(n,n)=\prod_{n}^{1}(1-k_i^2)^{-1/2},$$

$$\Phi(n,z=1)=\sum_{t=-n}^{n}\Phi(n,t)=\prod_{n}^{1}\left(\frac{1-k_i}{1+k_i}\right)^{1/2}, \tag{B.5}$$

$$\Phi(n,-n)=-k_n\prod_{n}^{1}(1-k_i^2)^{-1/2}.$$

Note also that we have

$$\Psi(n,t)+\Psi(n,-t)=\Phi(n,t)+\Phi(n,t)=\sum_{i,j}m_{ij}(n,t). \tag{B.6}$$

Acknowledgments. It is a pleasure to thank Dr. Hanoch Lev-Ari of Stanford, and Prof. Bernard C. Levy of the Massachusetts Institute of Technology, Cambridge, Massachusetts, for many interesting and useful discussions.

REFERENCES

[1] I. M. GELFAND AND B. M. LEVITAN, *On the determination of a differential equation from its spectral function*, Amer. Math. Soc. Transl., 1 (1955), pp. 253–304.

[2] Z. S. AGRANOVICH AND V. A. MARCHENKO, *The Inverse Problem of Scattering Theory*, Gordon and Breach, New York, 1963.

[3] M. G. KREIN, *On a method for the effective solution of the inverse boundary value problem*, Dokl. Akad. Nauk SSSR, 94 (1954), pp. 987–990 (in Russian).

[4] L. D. FADDEEV, *The inverse problem of quantum theory of scattering*, J. Math. Phys., 4 (1963), pp. 72–104.

[5] K. CHADAN AND P. C. SABATIER, *Inverse Problems in Quantum Scattering Theory*, Springer–Verlag, New York, 1977.

[6] K. M. CASE AND M. KAC, *A discrete version of the inverse scattering problem*, J. Math. Phys., 14 (1973), pp. 594–603.

[7] K. M. CASE, *The discrete inverse scattering problem on one dimension*, J. Math. Phys., 15 (1974), pp. 143–146.

[8] C. L. PEKERIS, *Direct method of interpretation in resistivity prospecting*, Geophysics, 5 (1940), pp. 31–42.

[9] E. A. ROBINSON, *Dynamic predictive deconvolution*, Geophys. Prosp., 23 (1975), pp. 779–797.

[10] V. BARDAN, *Comments on dynamic predictive deconvolution*, Geophys. Prosp., 25 (1977), pp. 569–572.

[11] E. A. ROBINSON AND S. TREITEL, *Geophysical Signal Processing*, Prentice–Hall, Englewood Cliffs, NJ, 1980, pp. 492–505.

[12] P. HUBRAL, *On getting reflection coefficients from waves*, Geophys. Prosp., 26 (1978), pp. 627–630.

[13] E. A. ROBINSON, *Spectral approach to geophysical inversion by Lorentz, Fourier and radon transforms*, Proc. IEEE, 70 (1982), pp. 1039–1054.

[14] G. CYBENKO, *The numerical stability of the Levinson–Durbin algorithm for Toeplitz systems of equations*, SIAM J. Sci. Statist. Comput., 1 (1980), pp. 303–319.

[15] A. BULTHEEL, *Error analysis of incoming and outgoing schemes for the trigonometric moment problem*, Proc. Conference on Rational Approximation, Amsterdam, Springer–Verlag, New York, 1981.

[16] J. A. WARE AND K. AKI, *Continuous and discrete inverse scattering problems in a stratified elastic medium* I: *plane waves at normal incidence*, J. Acoust. Soc. Amer., 45 (1969), pp. 911–921.

[17] J. G. BERRYMAN AND R. R. GREENE, *Discrete inverse methods for elastic waves in layered media*, Geophysics, 45 (1980), pp. 213–233.

[18] J. M. MENDEL AND F. HABIBI-ASHRAFI, *A survey of approaches to solving inverse problems for lossless layered media systems*, IEEE Trans. Geosci. Remote Sensing, GE-18 (1980), pp. 320–330.

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

[19] R. G. NEWTON, *Inversion of reflection data for layered media: A review of exact methods*, Geophys. J. Royal Astr. Soc., 65 (1981), pp. 191–215.

[20] R. BURRIDGE, *The Gelfand–Levitan, the Marchenko and the Gopinath–Sondhi integral equations of inverse scattering theory regarded in the context of inverse impulse response problems*, Wave Motion, 2 (1980), pp. 305–323.

[21] P. DEIFT AND E. TRUBOWITZ, *Inverse scattering on the line*, Comm. Pure Appl. Math., 32 (1979), pp. 121–251.

[22] W. SYMES, *Stable solution of the inverse reflection problem for a smoothly stratified medium*, SIAM J. Math. Anal., 12 (1981), pp. 421–453; see also *Impedance profile inversion via the first transport equation*, J. Math. Anal. Appl., 94 (1983), pp. 435–453.

[23] F. SANTOSA AND H. SCHWETLICK, *The inversion of acoustical impedance profiles by methods of characteristics*, Wave Motion, 4 (1982), pp. 99–110.

[24] M. M. SONDHI AND J. R. RESNICK, *The inverse problem for the vocal tract: numerical methods, acoustical experiments and speech synthesis*, J. Acoust. Soc. Amer., 73 (1983), pp. 985–1002.

[25] A. M. BRUCKSTEIN, B. C. LEVY AND T. KAILATH, *Differential methods in inverse scattering*, ISL Report, Dept. of Electrical Engineering, Stanford Univ., Stanford, CA, 1983; SIAM J. Appl. Math, 45 (1985), pp. 312–335.

[26] A YAGLE AND B. C. LEVY, *The Schur algorithm and its applications*, LIDS Report, Massachusetts Inst. of Technology, 1983; Acta Appl. Math., 3 (1985), pp. 255–284.

[27] K. P. BUBE AND R. BURRIDGE, *The one-dimensional inverse problem of reflection seismology*, this Journal, 25 (1983), pp. 497–559.

[28] A. M. BRUCKSTEIN, I. KOLTRACHT AND T. KAILATH, *Inverse scattering with noisy data*, preprint, ISL, Stanford Univ., 1984; SIAM J. Sci. Statist. Comput., 7 (1986), pp. 1331–1349.

[29] G. ARSENE, Z. CEAUSESCU, F. POTRA AND D. TIMOTIN, *A new algorithm for detecting reflection coefficients in layered media*, Ann. Geophysicae, 1, 4–5 (1983), pp. 285–290.

[30] T. K. CITRON, A. M. BRUCKSTEIN AND T. KAILATH, *An inverse scattering interpretation of the partial realization problem*, Proc. 23rd CDC Conference, Las Vegas, 1984.

[31] T. K. CITRON AND T. KAILATH, *Euclid's algorithm, scattering theory and a* VLSI *architecture for decoding Reed–Solomon codes*, IEEE Trans. Inform. Theory, submitted for publication, 1985.

[32] E. HAIRER, CH. LUBICH AND M. SCHLICHTE, *Fast numerical solution of nonlinear Volterra convolution equations*, SIAM J. Sci. Statist. Comput., 6 (1985), pp. 532–541.

[33] I. SCHUR, *Über Potentzreihen, die im Innern des Einheitskreises Beschrankt sind*, J. Reine Angew. Math., 147 (1917), pp. 205–232.

[34] N. I. AKHIEZER, *The Classical Moment Problem*, Hafner, New York, 1965.

[35] P. DEWILDE, A. C. VIEIRA AND T. KAILATH, *On a generalized Szego–Levinson realization algorithm for optimal linear predictors based on a network synthesis approach*, IEEE Trans. Circuits and Systems, CAS-25 (1978), pp. 663–675.

[36] P. DEWILDE AND H. DYM, *Schur recursions, error formulas and convergence of rational estimators for stationary stochastic processes*, IEEE Trans. Inform. Theory, IT-27 (1981), pp. 446–461.

[37] ———, *Lossless chain scattering matrices and optimum linear prediction*, Circuit Theory Appl., 9 (1981), pp. 135–175.

[38] P. DEWILDE, J. T. FOKKEMA AND I. WIDYA, *Inverse scattering and linear prediction, the time continuous case*, in Stochastic Systems: Mathematics of Filtering and Identification Applications, D. Reidel, Dordrecht, The Netherlands, 1981.

[39] S. K. RAO AND T. KAILATH, *Orthogonal digital filters for* VLSI *implementation*, IEEE Trans. Circuits and Systems, CAS-31 (1984), pp. 993–945; see also VLSI *arrays for digital signal processing, Part* I: *A model identification approach to digital filter realizations*, IEEE Trans. Circuits and Systems, 32 (1985), pp. 1105–1118.

[40] B. D. O. ANDERSON AND T. KAILATH, *Fast algorithms for the integral equations of the inverse scattering problem*, Integral Equations Operator Theory, 1/1 (1978), pp. 132–136.

[41] H. DYM AND H. P. MCKEAN, *Gaussian Processes, Function Theory and the Inverse Scattering Problem*, Academic Press, New York, 1976.

[42] H. J. LANDAU, *The inverse problem for the vocal tract and the moment problem*, SIAM J. Appl. Math., 14 (1983), pp. 1019–1035.

[43] A. E. YAGLE AND B. C. LEVY, *Layer-stripping solutions of multi-dimensional inverse scattering problems*, LIDS Report LIDS-P-1475, Electrical Engineering Dept., Massachusetts Inst. of Technology, 1985; also J. Math. Phys., 27 (1986), pp. 1701–1710.

[44] L. WOHLERS, *Lumped and Distributed Passive Networks*, Academic Press, New York, 1969.

[45] B. GOPINATH AND M. M. SONDHI, *Inversion of the telegraph equation and the synthesis of nonuniform lines*, Proc. IEEE, 59 (1971), pp. 383–392.

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

[46] R. REDHEFFER, *On the relation of transmission-line theory to scattering and transfer*, J. Math. Phys., 41 (1962), pp. 1–41.

[47] T. KAILATH, A. M. BRUCKSTEIN AND D. MORGAN, *Fast matrix factorizations via discrete transmission lines*, Linear Algebra Appl., 75 (1986), pp. 1–25.

[48] T. KAILATH AND H. LEV-ARI, *On mappings between covariance matrices and physical systems,* Proc. AMS Summer Conference, Bowdoin College, Brunswick, ME, 1984.

[49] W. W. SYMES AND G. ZIMMERMAN, *Experiments in impedance profile inversion using noisy and bandlimited data*, AMOCO Research Report, 1982.

[50] J. P. CORONES, M. E. DAVISON AND R. J. KRUEGER, *Direct and inverse scattering in the time domain via invariant imbedding equations*, J. Acoust. Soc. Amer., 74 (1983), pp. 1535–1541.

[51] T. KAILATH, *Time-variant and time-invariant lattice filters for nonstationary processes*, in Mathematical Tools and Models for Control Systems Analysis and Signal Processing, Vol. 2, Editions CNRS, Paris, 1982, pp. 417–464.

[52] H. LEV-ARI AND T. KAILATH, *Lattice filter parametrizations and modeling of nonstationary processes*, IEEE Trans. Inform. Theory, IT-30 (1984), pp. 2–16.

[53] J. M. DELOSME AND M. MORF, *A tree classification of algorithms for Toeplitz and related equations including generalized Levinson and doubling type algorithms*, Proc. 19th C.D.C., Albuquerque, 1980.

[54] D. MORGAN, *Orthogonal triangularization methods for fast estimation algorithms*, Ph.D. thesis, Dept. Electrical Engineering, Stanford Univ., 1985.

[55] B. R. MUSICUS, *Levinson and fast Cholesky algorithms for Toeplitz and almost Toeplitz matrices*, Internal Report, Electrical Engineering Dept., Massachusetts Inst. of Technology, November 1981.

[56] T. KAILATH, L. LJUNG AND M. MORF, *Generalized Krein–Levinson equations for efficient calculation of Fredholm resolvents of nondisplacement kernels*, in Topics in Functional Analysis, essays in honor of M. G. Krein, Advances in Mathematics Supplementary Studies, Vol. 3, Academic Press, New York, 1978, pp. 169–184.

[57] T. KAILATH, B. C. LEVY, L. LJUNG AND M. MORF, *The factorization and representation of operators in the algebra generated by Toeplitz operators*, SIAM J. Appl. Math., 37 (1979), pp. 467–484.

[58] M. MORF, *Fast algorithms for multivariable systems*, Ph.D. thesis, Electrical Engineering Dept., Stanford Univ., 1974.

[59] J. RISSANEN, *Algorithms for triangular decomposition of block Hankel and Toeplitz matrices with applications to factoring positive matrix polynomials*, Math. Comp., 27 (1973), pp. 147–154.

[60] I. GOHBERG, T. KAILATH AND I. KOLTRACHT, *Fast matrix factorization*, ISL Technical Report, Electrical Engineering Dept., Stanford Univ., 1986; see also *Efficient solution of linear systems of equations with recursive structure*, Linear Algebra Appl., 80 (1986), pp. 81–113.

[61] S Y. KUNG AND H. HU, *A highly concurrent algorithm and pipelined architecture for solving Toeplitz systems*, IEEE Trans. Acoust. Speech Signal Process., ASSP-31/1 (1983), pp. 66–76.

[62] I. GOHBERG AND M. G. KREIN, *Theory and applications of Volterra operators in Hilbert space*, Trans. Math. Monographs, 24, American Mathematical Society, Providence, 1970.

[63] R. CAFLISCH, *An inverse problem for Toeplitz matrices and the synthesis of discrete transmission lines*, Linear Algebra Appl., 38 (1981), pp. 207–225.

[64] A. M. BRUCKSTEIN AND T. KAILATH, *Some matrix factorization identities in inverse scattering*, Linear Algebra Appl., 74 (1986), pp. 157–172.

[65] ———, *Spatio-temporal scattering and inverse problems*, ISL Technical Report, Electrical Engineering Dept., Stanford Univ., 1983.

Downloaded 01/09/18 to 132.68.36.165. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php