Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

J.W. van der Woude

System identification of lossless layered media from input-output data

---

Department of Operations Research, Statistics, and System Theory Report BS-R8929 November

---

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

System Identification of Lossless Layered Media from Input-Output Data

J. W. van der Woude
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

In this paper we model the interior of the earth as a lossless layered medium, and we are concerned with the reconstruction of the parameters of such a medium from the input and the output (input-output data). In the first part of the paper we assume that the input-output data is noise free, and we derive a so-called layer stripping algorithm for the reconstruction of the parameters of the medium. The algorithm is just one of the many versions of layer stripping algorithms that are available in the literature. In the second and major part of the paper we assume that both the input and the output are corrupted by noise, and we derive two new methods for the estimation of the parameters of the medium using the noisy input-output data. The methods are based on the estimation of the parameters of an ARX-representation of a lossless layered medium. The distinction between the two methods is that in one method the ARX parameters are free, while in the other method the ARX parameters have to correspond to parameters of a lossless layered medium that are physically realistic. We illustrate the methods by means of a numerical experiment.

1980 Mathematics Subject Classification (1985): 73D50, 86A15, 93E12.

Key Words and Phrases: Lossless layered medium,reflection coefficients, parameter estimation, input-output data, ARX-representation. problem.

1. Introduction

In the present paper we model the interior of the earth as a system of horizontal layers through which signals only travel in the vertical direction. We assume that each layer is lossless and that each layer has a thickness such that a signal needs one unit of time to pass the layer. When a signal arrives at the boundary of two layers with distinct propagation velocities, part of the signal is reflected and part is transmitted. The rate of reflection and transmission at a boundary is determined by the reflection coefficient of the boundary layer. Amongst others, this coefficient is determined by the propagation velocities in the two adjacent layers, which in turn give an indication of the physical properties of the two layers. For oil prospecting purposes it is useful to know the physical properties of the layers. Therefore, there is a need for methods by which we can determine the reflection coefficients of the boundary layers. However, we can not directly determine or measure the reflection coefficients. Things that can be measured are for instance signals that come up from the earth. Such signals can be artificially generated by sending a signal, in this paper called a *source wavelet*, from the surface into the earth. Then, to determine the reflection coefficients, we have to solve a so-called inverse problem, which consists of the reconstruction of the reflection coefficients of a lossless layered medium from the source wavelet and the *seismogram*, the measurement of the reflected signal.

In this paper we distinguish two situations. In the first part of the paper we assume that the source wavelet and the seismogram are free of noise. Then, based on the fact that we modelled the interior of the earth as a lossless layered medium, we derive a method for the reconstruction of the reflection coefficients from the source wavelet and the seismogram. The obtained method is just one of the many

versions that are available in the literature of so-called layer stripping methods. It turns out that in the method we do not need a-priori knowledge about the number of layers. By means of an example we show that the method performs well in case the source wavelet and the seismogram are not corrupted by noise. We also show that the method may come up with poor results when the source wavelet and the seismogram are corrupted by noise. The method may then produce reflection coefficients that are not physically realistic.

In the second and major part of this paper we assume that we only have available noise corrupted recordings of the source wavelet and the seismogram. Then, based on a so-called ARX-representation of a lossless layered medium, we develop two new methods for the reconstruction of the reflection coefficients from the noisy source wavelet and seismogram. The first of these two methods can be considered to be an extension of the layer stripping method mentioned before, because in the method the data is smoothed first in the sense that a set of ARX-parameters is estimated, and then the smoothed data are processed in a layer stripping like fashion. A drawback of the method is that it may come up with reflection coefficients that are not physically realistic. The second of the two methods is a refinement of the first one and does not have the drawback of producing unrealistic reflection coefficients. This is caused by the fact that in the method the ARX-parameters to be estimated are forced to correspond to physically realistic reflection coefficents. For both methods, in contrast to layer peeling methods, we need a-priori knowledge about the number of layers.

The outline of the present paper is as follows. In section 2 we give a description of the interior of the earth modelled as a lossless layered medium. The model is parametrized by the reflection coefficients. Also in section 2 we describe how vertical travelling signals are scattered at the boundaries between layers. In section 3 we present the main features of an algorithm to compute a so-called synthetic seismogram, given a source wavelet and the reflection coefficients. In section 4 we present a layer stripping method to solve the inverse of the problem of section 3. This inverse problem consists of the reconstruction of the reflection coefficients from the source wavelet and the seismogram. In section 5 and 6 we develop the two methods, mentioned before, in which the presence of noise in the source wavelet and the seismogram is explicitely taken into account. In section 7 we present some of the computational aspects of the methods. It turns out that in the computations we can make use of the efficient Levinson recursions, which are well-known from the theory of signal processing. In section 8 we present the results of some numerical experiments done with the proposed algorithms. In section 9 we offer some remarks and conclusions.

We believe that in the context of seismic signal processing the approach described in the second part of this paper is new. In all recent literature on the reconstruction of reflection coefficients from noisy data, it is always assumed that only noise is present in the seismogram, and that the source wavelet is known exactly. The reason that we also allow uncertainty in the source wavelet is that in practical situations the source wavelet is often only known approximately. In fact, frequently the source wavelet is not measured at all, but is the result of some rules of thumb or is taken from standard tables. Two additional nice features of the approach in the second part of the paper are the fact that in the reconstruction all data are used to estimate any reflection coefficient, and the fact that it is not required to know at which time the signals actually start.

The investigations reported in this paper are supported by a fellowship from Shell Research B.V.

2. Model Description

In this section we describe how the interior of the earth can be modelled as a lossless layered medium. Therefore, we first consider the scattering of signals at a boundary between two layers.

2.1. Reflections and transmissions at a boundary

In this subsection we consider two adjacent homogeneous lossless layers. We assume that through each of the two layers a signal with a unit amplitude travels along the vertical in the direction of the common boundary. Then, if the propagation velocities in the two layers are different, on arrival at the boundary, part of each signal is reflected and part is transmitted. The signals are said to be *scattered.* The scattering of the two signals can be depicted as in the figures 1.a and 1.b, respectively. In these figures the horizontal axis corresponds with time, while the vertical axis corresponds with depth.

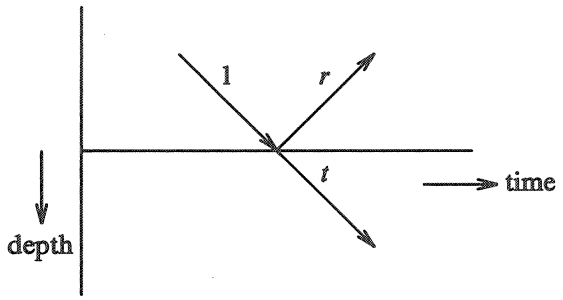

Fig 1.a. Scattering of downgoing signal

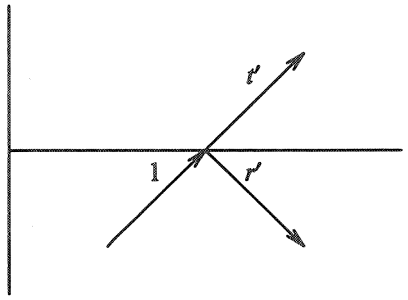

Fig 1.b. Scattering of upgoing signal

In the figures 1.a and 1.b, the amplitudes of the reflected signals are denoted r and r', and the amplitudes of the transmitted signals are denoted t and t'. The amplitudes of the scattered signals satisfy $t=1+r$, $t'=1+r'$, $r=-r'$ and $|r|\leqslant 1$, where the fact is used that the layers are lossless (cf. Claerbout [5]). Hence, the scattering of the signals at the boundary can be expressed by a single parameter r, called the reflection coefficient of the boundary (with respect to downgoing signals).

If two vertically traveling signals in two adjacent layers arrive at the common boundary we have a situation as depicted in figure 2.

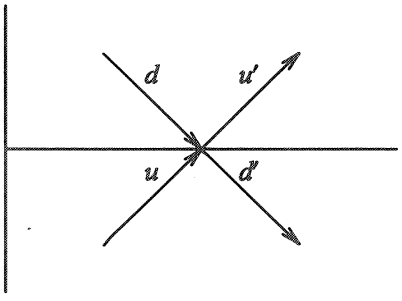

Fig 2. Scattering at boundary

In figure 2 d and u represent the amplitudes of the downgoing and upgoing signals before scattering, while d' and u' represent the amplitudes of the downgoing and the upgoing signals after scattering. Using superposition it follows that $d' = t\,d+r'\,u = (1+r)\,d-r\,u$ and $u' = r\,d+t'\,u = r\,d+(1-r)\,u$, or in matrix/vector notation

$$\begin{bmatrix} d' \\ u' \end{bmatrix} = \begin{bmatrix} 1+r & -r \\ r & 1-r \end{bmatrix} \begin{bmatrix} d \\ u \end{bmatrix}. \tag{1}$$

Also it follows that

$$\begin{bmatrix} d' \\ u \end{bmatrix} = \frac{1}{(1-r)} \begin{bmatrix} 1 & -r \\ -r & 1 \end{bmatrix} \begin{bmatrix} d \\ u' \end{bmatrix} \tag{2}$$

provided that $1-r \neq 0$. If $r = 1$ then the boundary is a perfect reflector for downgoing signals (cf. Claerbout [5]). These perfect reflectors for both downgoing and upgoing signals ($r = -1$) do not occur in practice and therefore we assume in the sequel that all reflection coefficients have absolute value less than one ($|r| < 1$).

2.2. Multiple layers

In this subsection we assume that the interior of the earth consists of a system of k layers located on top of a basement. This means that there are $k + 1$ boundaries. We assume that the layers have a thickness such that a signal needs one unit of time to travel through any of the k layers. We also assume that before time $t = 0$ no signals are present in the system of layers. Furthermore, we assume that from $t = 0$ on, from the top boundary, a source wavelet is sent into the earth along the vertical, and that, for all time t, no signals come up from the basement. Finally, we denote m for the source wavelet that is sent into the earth, and y for the seismogram, the signal that is reflected by the earth and that passes the top boundary. The process of scattering at time t in all layers can then be depicted as in figure 3.

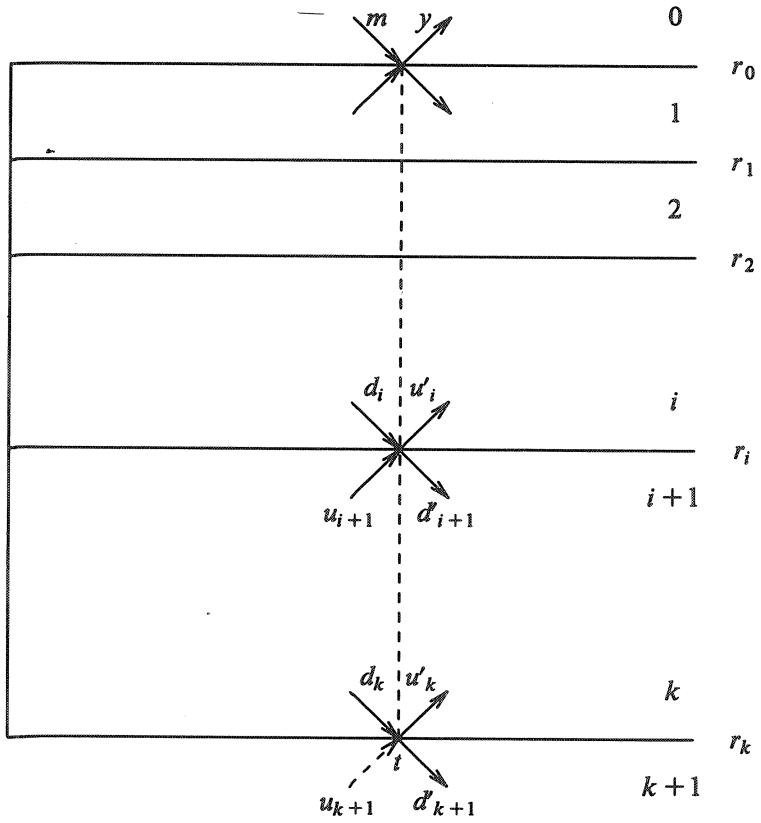

Fig 3. Scattering at the boundaries at time t

In figure 3 the reflection coefficient at the boundary between layer i and layer $i+1$ is denoted r_i. The downgoing and upgoing signals before scattering in layer i are denoted d_i and u_i, and are defined at the bottom and the top of layer i, respectively. The downgoing and upgoing signals after scattering in layer i are denoted d'_i and u'_i, and are defined at the top and the bottom of layer i, respectively. Because we

assume that no signals come up from the basement, it follows that $u_{k+1}(t) = 0$ for all t. Furthermore, if we consider the medium above the top layer to be layer 0, we can denote $m(t) = d_0(t)$ and $y(t) = u'_0(t)$ for all t. Since we assume that each of the k layers is travelled without loss of energy in one unit of time, it follows that $d_i(t+1) = d'_i(t)$ and $u_i(t+1) = u'_i(t)$ for all t and for all $i = 1,2,...,k$. Combining the above, we obtain

$$\begin{bmatrix} d_{i+1}(t+1) \\ u_i(t+1) \end{bmatrix} = \begin{bmatrix} 1+r_i & -r_i \\ r_i & 1-r_i \end{bmatrix} \begin{bmatrix} d_i(t) \\ u_{i+1}(t) \end{bmatrix} \tag{3}$$

for all t and for all $i = 0,1,...,k$, with $d_0(t) = m(t)$, $u_0(t+1) = y(t)$ and $u_{k+1}(t) = 0$ for all t.

As noted before, we assume that before $t = 0$ no signals are present in the system of layers. This means that $d_i(t) = 0$ and $u_i(t) = 0$ for all $t < 0$ and $i = 0,1,...,k$. From this it follows by (3) that $d_i(t) = 0$ for all $t < i$ and $u_i(t) = 0$ for all $t < i+1$ and $i = 0,1,...,k$, because it takes i units of time for a downgoing signal to travel from the surface layer to the i^{th} layer, and its contribution to the upgoing signal in the i^{th} layer even takes one unit of time more. Furthermore, from (3) it follows that

$$\begin{bmatrix} d_{i+1}(t+1) \\ u_{i+1}(t) \end{bmatrix} = \frac{1}{(1-r_i)} \begin{bmatrix} 1 & -r_i \\ -r_i & 1 \end{bmatrix} \begin{bmatrix} d_i(t) \\ u_i(t+1) \end{bmatrix}. \tag{4}$$

for all t and $i = 0,1,...,k$, because we assumed that $|r_i| < 1$ for all $i = 0,1,...,k$.

Note that all equations until sofar only establish relations between functions. They are neither differential nor difference equations. Since it is the ultimate goal to apply the results to real-life data which in general is sampled data, it seems natural to take the unit of time, here chosen one, equal to a multiple of the sampling time and to consider all equations to be 'discrete time' difference equations. Then using these difference equations we can calculate the seismogram y resulting from the source wavelet m and the reflection coefficients r_i.

3. Scattering

In this section we describe the main features of an algorithm for the calculation of the seismogram y resulting from a source wavelet m using the difference equations (3). The signal y obtained in this way is sometimes called a synthetic seismogram. The data for the computation of the signal y are the values of the source wavelet m at times $t = 0,1,...,T$, the number of layers k and the $k+1$ reflection coefficients $r_0, r_1, \ldots, r_k$.

To describe the main features of the algorithm, we denote $d(i) = d_i(t)$, $u(i) = u_i(t)$, $dd(i) = d_i(t+1)$ and $uu(i) = u_i(t+1)$ for all $i = 0,1,...,k+1$ at a given time t, $0 \leqslant t \leqslant T$. According to (3) the following relations are then satisfied,

$$\begin{aligned} &d(0) = m(t), \\ &dd(i+1) = (1+r_i)\,d(i) - r_i\,u(i+1) \\ &uu(i) \quad\;\; = r_i\,d(i) + (1-r_i)\,u(i+1) \end{aligned} \quad \text{for } i = 0,1,...,k-1,$$

($u_{k+1}(t) = 0$ for all t and $d_{k+1}(t)$ is of no interest).

$$uu(k) = r_k d(k),$$

$$y(t) = uu(0).$$

Then, using the following update rule

$$d(i) := dd(i),\ u(i) := uu(i) \ \text{ for } i = 1,2,...,k, \ \ t := t+1$$

the above relations provide us with an algorithm to calculate the seismogram $y(t)$ for $t, 0 \leqslant t \leqslant T$, given the reflection coefficients r_i, $i = 0,1,...,k$ and given the source wavelet $m(t)$ for $t, 0 \leqslant t \leqslant T$. The algorithm can be initialized by setting $d(i) = 0$ and $u(i) = 0$ for $i = 0,1,...,k+1$ at time $t = 0$.

A numerical experiment with the obtained algorithm is presented in example 8.1 in section 8.

4. Inverse Scattering - the noise free case

In this section we present a method for solving the inverse of the problem of the previous section. This inverse problem can be formulated as follows. Given the source wavelet m and the seismogram y generated by the equations (3), find the underlying reflection coefficients. The method for calculating the coefficients is derived using equations (3) and (4), and is based on the following observation (see also the remarks below (3)). Let v be the smallest integer such that $m(v) \neq 0$, i.e. $v = \min \{t, 0 \leqslant t \leqslant T \mid m(t) \neq 0\}$ is the time that the source wavelet m 'starts'. Then it follows from (3) that $d_i(v+i) \neq 0$ and $u_i(v+i) = 0$ for all $i = 0,1,...,k$. This means that $u_i(v+i+1) - r_i d_i(v+i) = 0$ for $i = 0,1,...,k$, or

$$r_i = \frac{u_i(v+i+1)}{d_i(v+i)} \quad \text{for } i = 0,1,...,k.$$

For instance, this implies that $r_0 = \frac{y(v)}{m(v)}$.

Using the above observations we can now describe the main features of an algorithm for solving the inverse problem. Therefore, for a given $i \geqslant 0$, representing the number of a layer, we denote $d(t) = d_i(t+v)$, $u(t) = u_i(t+v+1)$, $dd(t) = d_{i+1}(t+v)$ and $uu(t) = u_{i+1}(t+v+1)$ for $t = i, i+1,...,T-v-i$. The signals d and u can be thought of as the downgoing and the upgoing signals, respectively, at the bottom of layer i, while dd and uu can be thought of as the downgoing and upgoing signals, respectively, at the bottom of layer $i+1$. Then it follows from (3) and (4) that

$$r_i = \frac{u(i)}{d(i)},$$

$$dd(t+1) = \frac{1}{1-r_i}(d(t) - r_i u(t)) \quad \text{for } t = i, i+1,...,T-v-i-1,$$

$$uu(t-1) = \frac{1}{1-r_i}(u(t) - r_i d(t)) \quad \text{for } t = i+1, i+2,...,T-v-i.$$

Now using the update rule

$$d(t) := dd(t),\ u(t) := uu(t) \text{ for } t = i+1,...,T-v-i-1,\ i := i+1$$

we obtain an algorithm to compute the reflection coefficients r_i for $i \geqslant 0$, from the source wavelet $m(t)$ and the seismogram $y(t)$ for t, $0 \leqslant t \leqslant T$. The algorithm can be initialized by setting $d(t) = m(t+v)$ and $u(t) = y(t+v)$ for $t = 0,1,..,T-v$ for $i = 0$.

The obtained algorithm is one of the many so-called layer stripping algorithms that can be found in the literature for solving inverse problems (cf. Bruckstein and Kailath [2], Bube and Burridge [4], and Ursin and Berteussen [10]). In example 8.2 in section 8 the algorithm is applied to the source wavelet $m(t)$ given, and the seismogram $y(t)$ calculated in example 8.1. In example 8.3 in section 8 the algorithm is applied to the data consisting of the data of example 8.2, corrupted by Gaussian white noise with zero mean and a standard deviation of 0.003. The results show that the method performs well in case the signals are free of noise, but that in case the signals are corrupted by noise the method may perform poorly. One of the reasons for the poor performance of the method is the fact that in the case of noisy signals the method can not determine exactly at which time the source wavelet actually 'starts'. This may lead to a situation, like in example 8.3, that only noise is used to determine the reflection coefficients. Another reason for a poor performance of the layer stripping method may be caused by the poor numerical properties of the method in the case that the reflection coefficients are 'large' (in absolute value close to 1) (cf. Bruckstein, Koltracht and Kailath [3]).

5. Inverse Scattering - the noisy case

5.1. Introduction

In the remainder of the present paper we assume that the input-output data is corrupted by noise and we present two methods in which this presence of noise is explicitly taken into account. However, before going into the details of the derivation, we want to stress that in the methods of the second part of this paper we assume that *both* the source wavelet and the seismogram are corrupted by noise. This in contrast with existing methods for the reconstruction of the reflection coefficients from noisy data (cf. Habibi-Ashrafi and Mendel [8], Goutsias and Mendel [9]). In the methods described in these references it is *always* assumed that the source wavelet is known exactly. However, as we mentioned before, this often is not the case. Often, as we assume in this paper, the source wavelet is only known approximately, and then the methods in the above references can not be applied.

5.2. ARX-representation

To develop both methods we make use of the transfer function from the source wavelet m to the seismogram y. For the introduction of this transfer function we make use of the Z-transform of a time series. Let $f = \{f(t) \mid t \geqslant 0\}$ be a time series, then the *Z-transform* of f, denoted $\hat{f}$, is formally defined as $\hat{f}(z) := \sum_{t \geqslant 0} f(t) z^{-t}$. If $g = \{f(t+1) \mid t \geqslant 0\}$ and $f(0) = 0$, then $\hat{g}(z) = z\hat{f}(z)$.

Now suppose that $R_i(z)$, for $i = 0,1,...,k$, denotes the transfer function from $d_i(t)$ to $u_i(t+1)$,i.e. $R_i(z) = z\hat{u}_i(z)/\hat{d}_i(z)$, where it is used that $u_i(0) = 0$ for $i = 0,1,...,k$. It follows from (3) that

$$\begin{bmatrix} z\hat{d}_{i+1}(z) \\ z\hat{u}_i(z) \end{bmatrix} = \begin{bmatrix} 1+r_i & -r_i \\ r_i & 1-r_i \end{bmatrix} \begin{bmatrix} \hat{d}_i(z) \\ \hat{u}_{i+1}(z) \end{bmatrix} \quad \text{for } i = 0,1,...,k-1.$$

Furthermore, since $u_{k+1}(t) = 0$ for all t, it is also clear from (3) that $R_k(z) = r_k$. Using $z\hat{u}_{i+1}(z) = R_{i+1}(z)\,\hat{d}_{i+1}(z)$ and $z\hat{u}_i(z) = R_i(z)\,\hat{d}_i(z)$, it can be shown that

$$R_i(z) = \frac{z^2 r_i + R_{i+1}(z)}{z^2 + r_i R_{i+1}(z)} \qquad \text{for } i = 0,1,...,k-1, \tag{5.a}$$

and

$$R_k(z) = r_k. \tag{5.b}$$

To compute the coefficients of $R_i(z)$, for $i = 0,1,...,k$, we write $R_i(z) = b_i(z)/a_i(z)$, where $a_i(z)$ and $b_i(z)$ are polynomials in z with real coefficients. Using (5.a) it then follows that

$$\frac{b_i(z)}{a_i(z)} = \frac{z^2 r_i a_{i+1}(z) + b_{i+1}(z)}{z^2 a_{i+1}(z) + r_i b_{i+1}(z)} \qquad \text{for } i = 0,1,...,k-1.$$

So, the polynomials $a_i(z)$ and $b_i(z)$ can be calculated by

$$\begin{aligned} a_i(z) &= z^2 a_{i+1}(z) + r_i b_{i+1}(z) \\ b_i(z) &= z^2 r_i a_{i+1}(z) + b_{i+1}(z) \end{aligned} \qquad \text{for } i = k-1,...1,0, \tag{6.a}$$

with

$$a_k(z) = 1\,,\; b_k(z) = r_k. \tag{6.b}$$

It can be shown that each polynomial $a_i(z)$ only has zeros in the open unit circle if $|r_j| < 1$ for $j = i, i+1,..., k$, and that each pair of polynomials $a_i(z)$ and $b_i(z)$ have no zeros in common if $|r_j| < 1$ for $j = i, i+1,..., k-1$ and $r_k \neq 0$. From (6) it is clear that both $a_i(z)$ and $b_i(z)$ only contain even powers of z. By induction it can be shown that the coefficient of $z^{2(k-i)}$ in $a_i(z)$, which is the coefficient of the leading power in $a_i(z)$, is equal to 1, and that the coefficient of $z^{2(k-i)}$ in $b_i(z)$ is equal to r_i. So, if $a_i(z)$ and $b_i(z)$ are computed according to (6), then r_i can be recovered as the coefficient of the leading power of $b_i(z)$, and also the polynomials $a_{i+1}(z)$ and $b_{i+1}(z)$ can be recovered from $a_i(z)$ and $b_i(z)$.

Indeed from (6) it follows that

$$a_{i+1}(z) = \frac{1}{z^2(1-r_i^2)}(a_i(z)-r_i b_i(z))$$
$$b_{i+1}(z) = \frac{1}{(1-r_i^2)}(b_i(z)-r_i a_i(z)) \qquad \text{for } i = 0,1,...k-1. \tag{7}$$

That $a_{i+1}(z)$ again is a polynomial is caused by the fact that the constant term in $a_i(z)-r_i b_i(z)$ is equal to zero. This can be proved by induction using (6). In summary, once $a_0(z)$ and $b_0(z)$ are known and are generated by (6), we can recover the underlying reflection coefficients r_i by computing and examining $a_i(z)$ and $b_i(z)$, for $i = 0,1,...,k$. Observe that only if the polynomials $a_0(z)$ and $b_0(z)$ are generated by equations of the type (6) then $a_i(z)$ and $b_i(z)$ computed via (7) will satisfy $a_k(z) = 1$ and $b_k(z) = r_k$. However, for polynomials of the same form as $a_0(z)$ and $b_0(z)$ but with arbitrary coefficients, $a_k(z)$ and $b_k(z)$, computed via (7), will not be constants, but will be proper rational functions and the recovered reflection coefficients may not be physically realistic.

Recall that $R_0(z) = z\hat{u}_0(z)/\hat{d}_0(z) = \hat{y}(z)/\hat{m}(z) = b_0(z)/a_0(z)$. So, $a_0(z)\hat{y}(z) = b_0(z)\hat{m}(z)$. If $a_0(z) = \sum_{i=0}^{k}\tilde{a}_i z^{2i}$ and $b_0(z) = \sum_{i=0}^{k}\tilde{b}_i z^{2i}$ this means that in the time domain

$$\sum_{i=0}^{k}\tilde{a}_i y(t+2i) = \sum_{i=0}^{k}\tilde{b}_i m(t+2i) \tag{8}$$

for all $t \geq 0$, where it is assumed that $m(t) = 0$ and $y(t) = 0$ for all $t = 0,1,...,2k-1$.

Note that in (8) only data points are related that are an even number of units of time apart from each other. This implies that the set of signals (m,y) that satisfy (8) for all $t \geq 0$ can be split into two sets of signals (m_1,y_1) and (m_2,y_2) that each satisfy (8) for all $t \geq 0$. Indeed, define $m_1(t) = m(t), y_1(t) = y(t)$, $m_2(t) = 0$ and $y_2(t) = 0$ for all even $t \geq 0$, and $m_2(t) = m(t), y_2(t) = y(t)$, $m_1(t) = 0$ and $y_1(t) = 0$ for all uneven $t \geq 0$. Then it is clear that both (m_1,y_1) and (m_2,y_2) satisfy (8) for all $t \geq 0$.

To simplify the discussion in the sequel, we introduce a new unit of time which is two times the old unit of time, and we only use the set (m_1,y_1), from now on denoted as (m,y). With these conventions (8) becomes

$$\sum_{i=0}^{k}\tilde{a}_i y(t+i) = \sum_{i=0}^{k}\tilde{b}_i m(t+i) \tag{9}$$

for all $t \geq 0$ where $m(t) = 0$ and $y(t) = 0$ for all $t = 0,1,...,k-1$. The relation (9) is a so-called autoregressive representation with exogeneous input (ARX representation) between the source wavelet m and the seismogram y.

6. Inverse scattering - a minimization problem

In this section we give a formulation for the problem of the reconstruction of reflection coefficients in the case that the data are corrupted by noise. Therefore, we assume that the source wavelet m and the seismogram y are known over a time interval such that the interval contains the supports of both signals and such that in the interval $m(t) = 0$ and $y(t) = 0$ for at least all $t < k$. The latter implies that we have to have an idea about the number of layers k. It also implies that we do not have to know exactly at which time the signals actually 'start'. We denote

$$M = (m(0),m(1),...,m(T))^T,\ Y = (y(0),y(1),...,y(T))^T,$$
$$Z = (Y^T,-M^T)^T,$$
$$A = (A_0^T,A_1^T,\ldots,A_{T-k}^T)^T,\ B = (B_0^T,B_1^T,\ldots,B_{T-k}^T)^T,$$

with

$$A_i = (0,0,...,0,\tilde{a}_0,\tilde{a}_1,\ldots,\tilde{a}_{k-1},\tilde{a}_k,0,...,0),$$
$$B_i = (0,0,...,0,\tilde{b}_0,\tilde{b}_1,\ldots,\tilde{b}_{k-1},\tilde{b}_k,0,...,0),$$

and

$$D = [A \; B].$$

Here $^{\mathsf{T}}$ denotes matrix transposition. A and B are matrices with rows A_i and B_i, respectively, having $T+1$ entries. Both A_i and B_i have i leading and $T-k-i$ trailling zero entries. D is a compound matrix made up of A and B. A, B and D are so-called Toeplitz matrices. Toeplitz matrices are matrices of which the entries along any sub- or super diagonal have the same value. Since $\tilde{a}_k = 1$ for all sets of reflection coefficients, it follows that the k^{th} super diagonal of A and D consists of entries that have the value 1. Hence, both A and D have full row rank for all sets of reflection coefficients. We denote $\theta = (\tilde{a}_0, \ldots, \tilde{a}_{k-1}, \tilde{b}_0, \ldots, \tilde{b}_k)$ for the vector of parameters of the ARX-representation (9), where we have omitted $\tilde{a}_k = 1$. Recall that the parameter vector θ can be determined from the vector of reflection coefficients $r = (r_0, r_1, \ldots, r_k)$ by (6). To express this dependency on r, we occasionally write $\theta(r)$. Furthermore, to express the dependency of D on θ, we write $D(\theta)$. It then follows that the ARX-representation (9) over the interval $[0, T]$ can also be described as

$$D(\theta)Z = 0 \tag{10}$$

If the signals m and y, contained in Z, are not corrupted by noise, then there exists a $\theta \in \mathbb{R}^{k+1}$ such that (9) and (10) hold exactly. In (9) and (10) it is still assumed that both signals m and y, contained in the vector Z, are known exactly. In that case the recovery of the vector of reflection coefficients r_i can be achieved by the layer stripping method described in the previous section, and the framework of ARX-representations is not necessary. However, in the remainder of the paper we assume that the input-output data is corrupted by noise, and it turns out that for the development of methods for the reconstruction of the reflection coefficients from the noisy data the framework of ARX-representations is very useful. Therefore, from now on we assume that only noisy recordings are available of the source wavelet and the seismogram. This means that the signals m and y are made up as follows

$$m(t) = \bar{m}(t) + m'(t), y(t) = \bar{y}(t) + y'(t)$$

for $t \geq 0$. Here $\bar{m}(t)$ and $\bar{y}(t)$ are the values of the noise free signals and $m'(t)$ and $y'(t)$ are the noise components at time t. The vector of true reflection coefficients is denoted $\bar{r} = (\bar{r}_0, \bar{r}_1, \ldots, \bar{r}_k)$ and the corresponding vector of true parameters is denoted $\bar{\theta}$. Clearly, we have that $\bar{\theta} = \theta(\bar{r})$. We denote

$$\bar{Z} = [\bar{y}(0), \bar{y}(1), \ldots, \bar{y}(T), -\bar{m}(0), -\bar{m}(1), \ldots, -\bar{m}(T)]^{\mathsf{T}},$$

$$Z' = [y'(0), y'(1), \ldots, y'(T), -m'(0), -m'(1), \ldots, -m'(T)]^{\mathsf{T}}$$

Then it follows from (10) that

$$D(\bar{\theta})\bar{Z} = 0.$$

However, the noise free vector $\bar{Z}$ is not available, but only the noisy vector Z. To reconstruct the reflection coefficients it is suggested to solve

$$D(\theta)Z = 0$$

for θ, and that a solution θ^* is a meaningful approximation of $\bar{\theta}$. In general however, the above equation for θ can *not* be solved exactly, and a natural thing to do is to apply a least squares approach to find a 'best possible solution' to the equation. This means that we have to solve the following problem.

PROBLEM 6.1.
Given Z determine $\hat{\theta}$ and $\hat{Z}$ such that $\| Z - \hat{Z} \|^2 = (Z - \hat{Z})^{\mathsf{T}}(Z - \hat{Z})$ is minimal, while $D(\hat{\theta})\hat{Z} = 0$.

To eliminate the constraint under which $\| Z - \hat{Z} \|^2$ is to be minimized, we apply the Lagrange multiplier method (cf. Ten Vregelaar [11]), and we therefore define

$$L(\hat{Z}, \hat{\theta}, \lambda) = \frac{1}{2}(Z - \hat{Z})^{\mathsf{T}}(Z - \hat{Z}) + \lambda^{\mathsf{T}} D(\hat{\theta})\hat{Z}.$$

Then taking the gradient with respect to $\hat{Z}$ and λ, we obtain $Z-\hat{Z} = D^{\mathsf{T}}(\hat{\theta})\lambda$ and $D(\hat{\theta})\hat{Z} = 0$. Because $D(\theta)$ has full row rank for all θ, it follows that $D(\theta)D^{\mathsf{T}}(\theta)$ is invertible for all θ. Consequently, $\lambda = (D(\theta)D^{\mathsf{T}}(\hat{\theta}))^{-1}D(\hat{\theta})Z$. Hence, $(Z-\hat{Z})^{\mathsf{T}}(Z-\hat{Z}) = J(\hat{\theta})$ where we have defined

$$J(\theta) = Z^{\mathsf{T}}D^{\mathsf{T}}(\theta)(D(\theta)D^{\mathsf{T}}(\theta))^{-1}D(\theta)Z.$$

The next problem is now equivalent to problem 6.1.

PROBLEM 6.2.
Given Z determine $\hat{\theta}$ such that $J(\hat{\theta})$ is minimal.

The equivalence of the two problems is obvious, because if $\hat{\theta}$ is known solving problem 6.2, then $\hat{\theta}$ together with $\hat{Z}$, defined as $\hat{Z} = (I - D^{\mathsf{T}}(\hat{\theta})(D(\hat{\theta})D^{\mathsf{T}}(\hat{\theta}))^{-1} D(\hat{\theta}))Z$, solves problem 6.1. Any $\hat{\theta}$ that with $\hat{Z}$ solves problem 6.1, obviously also solves problem 6.2.

When the noise components $m'(t)$ and $y'(t)$ are mutually independent zero mean white noise with variance σ^2, it has been shown by Aoki and Yue [1] that $\hat{\theta}$ solving the problem 6.2 yields a maximum likelihood estimate of $\bar{\theta}$. Furthermore, it has been shown by Ten Vregelaar [11] that in that case $P(\hat{\theta}\rightarrow\bar{\theta}$ as $T\rightarrow\infty) = 1$, provided the following conditions are satisfied:

1. The parameter space Θ is a known convex and compact subspace containing $\bar{\theta}$.
2. The source wavelet is bounded as $T\rightarrow\infty$.
3. The source wavelet is persistently exciting of order $2k+1$ (cf. [1]).
4. The polynomial $a(z) = \sum_{i=0}^{k-1}\tilde{a}_i z^i + z^k$ has only zeros in the open unit disk.
5. The polynomials $a(z)$ and $b(z) = \sum_{i=0}^{k}\tilde{b}_i z^i$ have no zeros in common.

Note that in the formulation of problem 6.2 it is not yet guaranteed that $\hat{\theta}$ is the parameter vector of an ARX-representation of a layered scattering system. In fact, this will be the distinction between the two methods presented in this section. In the first method the vector $\hat{\theta}$ may have any value, while in the second method the vector $\hat{\theta}$ may only have values that correspond to a vector $\hat{r}$ of physically realistic reflection coefficients. Hence, the first method comes down to solving the next problem.

PROBLEM 6.3.
Given Z determine $\hat{\theta}\in\mathbb{R}^{2k+1}$ such that $J(\hat{\theta})$ is minimal,

whereupon the reflection coefficients can be computed using (7).

For the second method, recall that if $\hat{\theta}$ is to be the parameter vector of an ARX-representation of a lossless layered medium, then $\hat{\theta}$ depends on a vector $\hat{r} \in (-1,+1)^{k+1}$ of physically realistic reflection coefficients. In that case $\hat{\theta} = \theta(\hat{r})$ and can be calculated from $\hat{r}$ by means of (6). The second method then comes down to solving the following problem.

PROBLEM 6.4.
Given Z determine $\hat{r}\in(-1,+1)^{k+1}$ such that $J(\theta(\hat{r}))$ is minimal.

7. COMPUTATIONAL ASPECTS

7.1. Introduction

In this section we discuss some of the computational aspects of the two methods described in the previous section. Common in both methods is that a function has to be minimized over some (sub)space. To find a vector for which the function has a (local) minimum, we propose the use of methods that, in addition to the value of the function, also require the gradient of the function.

7.2. The computation of the function $J(\theta)$

To compute $J(\theta)$ for a given $\theta \in \mathbb{R}^{2k+1}$ we have to solve the equation $D(\theta)D^{\mathsf{T}}(\theta)x = y$ with $y = D(\theta)Z$ for $x \in \mathbb{R}^{T-k+1}$. Denote $Q = D(\theta)D^{\mathsf{T}}(\theta)$, then because $D(\theta)$ is a full row rank matrix, it follows that Q is a symmetric positive definite matrix, i.e. $Q = Q^{\mathsf{T}}$ and $x^{\mathsf{T}}Qx > 0$ for all $x \neq 0$. From the special structure of $D(\theta)$, it also follows that Q is a Toeplitz matrix. So, Q is completely determined by its first column and we denote $q_i := q_{i0}$ for $i = 0,1,...,T-k$. Because Q is positive definite it follows that $q_0 > 0$, and by the special structure of $D(\theta)$ it follows from the Cauchy-Schwarz inequality that $q_0 > |q_i|$ for $i = 1,2,...,T-k$. Because of these properties, we can solve the equation $Qx = y$ by means of Levinson recursions (cf. Cybenko [6], Golub and Van Loan [7]), where it is shown that the numerical stability of these recursions is comparable with the numerical stability of solving $Qx = y$ by means of Choleski decompositions. The advantage of Levinson recursions over Choleski decompositions is that less computer memory is required (no matrices need to be stored) and the number of operations is of order kT, where k is the number of layers and T is the number of data. Hence, given any θ, $J(\theta)$ can be computed efficiently by successively computing $y = D(\theta)Z$, solving the equations $(D(\theta)D^{\mathsf{T}}(\theta))x = y$ by means of Levinson recursions, and computing $J(\theta) = x^{\mathsf{T}}y$.

In problem 6.4 it is required that θ depends on a vector of physically acceptable reflection coefficients r, expressed by writing $\theta(r)$. By means of (6) first $\theta(r)$ can be calculated from r, and then $J(\theta(r))$ can be computed as described above.

Both computation schemes are incorporated in two algorithms presented in subsection 7.4.

7.3. The computation of the gradient of $J(\theta)$

7.3.a. With respect to problem 6.3, note that

$$\frac{\partial}{\partial\theta_i}D(\theta) = [\frac{\partial}{\partial\theta_i}A(\theta) \;\; 0] \quad \text{and} \quad \frac{\partial}{\partial\theta_j}D(\theta) = [0 \;\; \frac{\partial}{\partial\theta_j}B(\theta)]$$

for $i = 0,1,...,k-1$ and for $j = k,k+1,...,2k$. The derivatives have a simple structure; it are matrices with entries 1 along a superdiagonal and with entries 0 elsewhere. Therefore, for $i = 0,1,...,2k$, $\frac{\partial}{\partial\theta_i}J(\theta)$, can be easily computed. In fact, if for the computation of $J(\theta)$ the vectors x and y are computed such that $y = D(\theta)Z$ and $D(\theta)D^{\mathsf{T}}(\theta)x = y$ then

$$\frac{\partial}{\partial\theta_i}J(\theta) = 2x^{\mathsf{T}}(\frac{\partial}{\partial\theta_i}D(\theta))(Z - D^{\mathsf{T}}(\theta)x),$$

for $i = 0,1,....,2k$.

7.3.b. The purpose of the second method is the estimation of physically acceptable reflection coefficients, and not the estimation of the parameters of the underlying ARX-representation. As noted before the vector θ depends on the vector r. Therefore, in the context of problem 6.4 it is more natural to work with the gradient of $J(\theta(r))$ with respect to r. The computation of this gradient requires the computation of $\frac{\partial}{\partial r_i}D(\theta(r))$ for $i = 0,1,...,k$, which comes down to the computation of $\frac{\partial}{\partial r_i}[\tilde{a}_0, \ldots, \tilde{a}_k]$ and $\frac{\partial}{\partial r_i}[\tilde{b}_0, \ldots, \tilde{b}_k]$ for $i = 0,1,...,k$. These computations, in turn, come down to the computation of $\frac{\partial}{\partial r_i}a(z)$ and $\frac{\partial}{\partial r_i}b(z)$ for $i = 0,1,...,k$. From (6) it follows that $a(z)$ and $b(z)$ can be calculated by the following recursions.

$$\begin{aligned} &a_k(z) = 1\,,\; b_k(z) = r_k. \\ &\left.\begin{aligned} a_j(z) &= za_{j+1}(z) + r_jb_{j+1}(z) \\ b_j(z) &= zr_ja_{j+1}(z) + b_{j+1}(z) \end{aligned}\right\} \quad \text{for } j = k-1,...,1,0, \\ &a(z) = a_0(z)\,,\; b(z) = b_0(z). \end{aligned}$$

Recall that at the end of section 5 we have doubled the sample time, causing the difference between the above recursions and the recursions described by (6). Now for a given $i, 0 \leqslant i \leqslant k$, denote $p_j(z) = \frac{\partial}{\partial r_i} a_j(z)$ and $q_j(z) = \frac{\partial}{\partial r_i} b_j(z)$ for all $j, 0 \leqslant j \leqslant i$. Then $p_0(z) = \frac{\partial}{\partial r_i} a(z)$ and $q_0(z) = \frac{\partial}{\partial r_i} b(z)$ can be computed as follows.

$$
\begin{aligned}
&p_i(z) = b_{i+1}(z),\ q_i(z) = z a_{i+1}(z), && \text{if } i < k,\\
&p_i(z) = 0,\ q_i(z) = 1, && \text{if } i = k,\\
&p_j(z) = z p_{j+1}(z) + r_j q_{j+1}(z),\\
&q_j(z) = z r_j p_{j+1}(z) + q_{j+1}(z), && \text{for } j = i-1, \ldots, 1, 0.
\end{aligned}
$$

By these recursions $\frac{\partial}{\partial r_i} D(\theta(r))$ can be computed for $i = 0, 1, \ldots, k$. Moreover, if the vectors x and y are computed such that $y = D(\theta(r))Z$ and $D(\theta(r))D^{\mathsf{T}}(\theta(r))x = y$ then

$$
\frac{\partial}{\partial r_i} J(\theta(r)) = 2x^{\mathsf{T}} \left(\frac{\partial}{\partial r_i} D(\theta(r))\right)(Z - D^{\mathsf{T}}(\theta(r))x)
$$

for $i = 0, 1, \ldots, k$.

7.3.c. To assure that the minimization process in the second method ends up with physically acceptable reflection coefficients ($\hat{r} \in (-1, +1)^{k+1}$), and to have the ability to apply unconstrained minimization routines, we have to introduce new variables. Therefore, we define the mapping $\phi: \mathbb{R}^{k+1} \to (-1, +1)^{k+1}$ as $\phi_i(s) = \frac{2}{\pi} \arctan(s_i)$, for $i = 0, 1, \ldots, k$. Then minimizing $J(\theta(r))$ under the constraint that $r \in (-1, +1)^{k+1}$, comes down to minimizing $J(\theta(\phi(s)))$ with respect to $s \in \mathbb{R}^{k+1}$.

7.4. Algorithms

The function $J(\theta)$ and the gradient of $J(\theta)$ for a given $\theta \in \mathbb{R}^{2k+1}$ as required in the first method can be calculated as follows.

ALGORITHM 7.1

1. Compute $y = D(\theta)Z$.
2. Compute x such that $D(\theta)D^{\mathsf{T}}(\theta)x = y$ by means of the Levinson recursion.
3. Compute $J(\theta) = x^{\mathsf{T}}y$.
4. For $i = 0, 1, \ldots, 2k$:
   Compute

$$
\frac{\partial}{\partial \theta_i} J(\theta) = 2x^{\mathsf{T}} \left(\frac{\partial}{\partial \theta_i} D(\theta)\right)(Z - D^{\mathsf{T}}(\theta)x)
$$

In the second method, $J(\theta(\phi(s)))$ and the gradient of $J(\theta(\phi(s)))$ for a given $s \in \mathbb{R}^{k+1}$ can be calculated as follows.

ALGORITHM 7.2

1. Compute $r = \phi(s)$.
2. Compute $\theta = (\tilde{a}_0, \ldots, \tilde{a}_{k-1}, \tilde{b}_0, \ldots, \tilde{b}_k)$ by means of the recursions to compute $a(z)$ and $b(z)$.
3. Compute $y = D(\theta)Z$.
4. Compute x such that $D(\theta)D^{\mathsf{T}}(\theta)x = y$ by means of the Levinson recursion.
5. Compute $J(\theta(\phi(s))) = x^{\mathsf{T}}y$.
6. For $i = 0, 1, \ldots, k$:
   a. Compute $\frac{\partial}{\partial r_i}\theta$ by means of the recursions to compute $\frac{\partial}{\partial r_i} a(z)$ and $\frac{\partial}{\partial r_i} b(z)$.

b. Compute

$$\frac{\partial}{\partial r_i} J(\theta(r)) = 2x^{\mathsf{T}}\left(\frac{\partial}{\partial r_i} D(\theta(r))\right)(Z - D^{\mathsf{T}}(\theta(r))x)$$

c. Compute $\frac{\partial}{\partial s_i} J(\theta(\phi(s))) = \frac{2}{\pi} \frac{1}{1+s_i^2} \frac{\partial}{\partial r_i} J(\theta(r))$.

Using a gradient method and the above schemes, we can find a solution to the following two problems.

PROBLEM 7.3
Given Z determine $\hat{\theta} \in \mathbb{R}^{2k+1}$ such that $J(\hat{\theta})$ is minimal.

PROBLEM 7.4
Given Z determine $\hat{s} \in \mathbb{R}^k$ such that $J(\theta(\phi(\hat{s})))$ is minimal.

If $\hat{\theta}$ is a solution of problem 7.3, then the underlying vector of reflection coefficients $\hat{r}$ can be calculated using (7). However, it is not guaranteed that $\hat{r}$ is a vector of physically realistic reflection coefficients. If $\hat{s}$ is a solution of problem 7.4, then the corresponding vector of reflection coefficients $\hat{r}$ can be calculated as $\hat{r} = \phi(\hat{s})$. Now it is guaranteed that $\hat{r}_i \in (-1, +1)$ for $i = 0, 1, ..., k$.

In examples 8.4 and 8.5 in section 8, the algorithms 7.1 and 7.2 in connection with a quasi Newton minimization method are applied to the noisy data of example 8.3 in order to reconstruct the reflection coefficients as given in example 8.1.

8. NUMERICAL EXPERIMENTS

In this section we illustrate by means of 5 examples of what is to be expected of the methods presented in this paper. In the examples we consider a lossless layered medium of 25 layers (26 boundaries) with reflection coefficients that are all zero, except $r_5 = 0.30$, $r_9 = -0.10$, $r_{16} = -0.27$, $r_{18} = 0.13$, $r_{25} = -0.21$. The source wavelet that we use to probe the system of layers is due to Habibi-Ashrafi and Mendel [8], and is given by

$$m(t) = 1360\, t\, e^{-500t} + 0.5\, e^{-15.3t} \sin\left(\frac{2\pi}{0.06} t\right).$$

For the layer stripping method as described in section 4, we have taken the unit of time to be 0.005, and we assume that the data are available over the time interval [0,1.2]. This means that for the two methods of the second part of the paper the unit of time is 0.01 and that T is 120. The reflection coefficients and the source wavelet are depicted in figure 4.

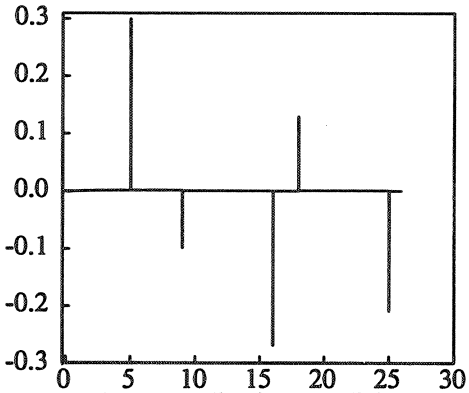

Fig 4.a. Reflection coefficients

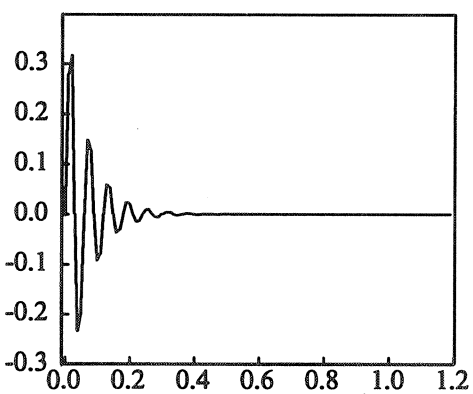

Fig 4.b. Source wavelet

EXAMPLE 8.1. Scattering.
In figure 5 the synthetic seismogram is depicted, computed from the source wavelet and the reflection coefficients using the algorithm described in section 3.

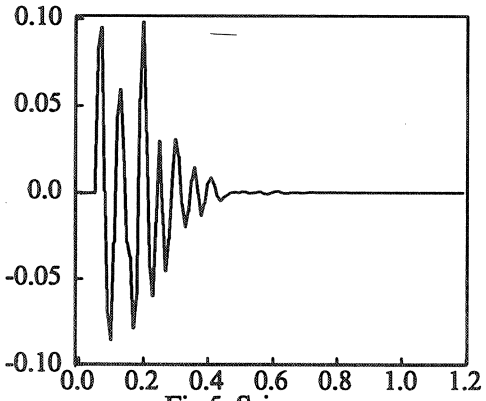

Fig 5. Seismogram

EXAMPLE 8.2. Layer stripping with noise free signals.
In this example the source wavelet and the seismogram of example 8.1 are used to reconstruct the reflection coefficients. The reconstruction is done by means of the layer stripping algorithm described in section 4. The first 25 reconstructed reflection coefficients are depicted in figure 6. From figures 3 and 6 it is clear that the layer stripping method performs well for noise free signals.

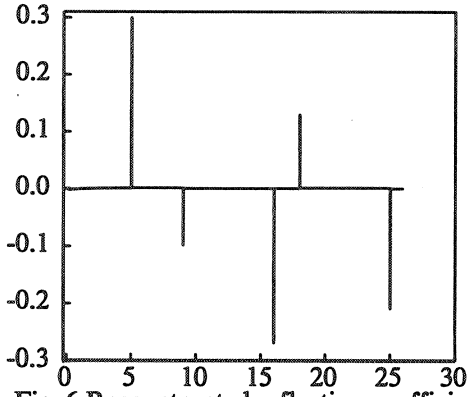

Fig. 6 Reconstructed reflection coefficients

EXAMPLE 8.3. Layer stripping with noisy signals.
In this example the layer stripping algorithm of example 8.2 is applied to the source wavelet and the seismogram of example 8.1, now both corrupted by Gaussian white noise with a zero mean and a standard deviation of 0.003. The noisy signals are depicted in figure 7, and the recovered reflection coefficients are depicted in figure 8. Again only the first 25 recovered reflection coefficients are depicted. From figure 7 it follows that in the case of noisy signals the layer stripping method may have a poor performance. The reason for this performance is that the time at which the source wavelet actually starts is not determined properly. In the present example the noise free source wavelet starts just after time $t = 0$, its value at $t = 0$ is zero, as is the value of the seismogram at $t = 0$, and these values are not used in the reconstruction. However, applying the method to the noisy signals, the source wavelet is considered to start at time $t = 0$, and the values of the source wavelet and the seismogram at $t = 0$, both noise, are used to determine the reflection coefficient of the top boundary.

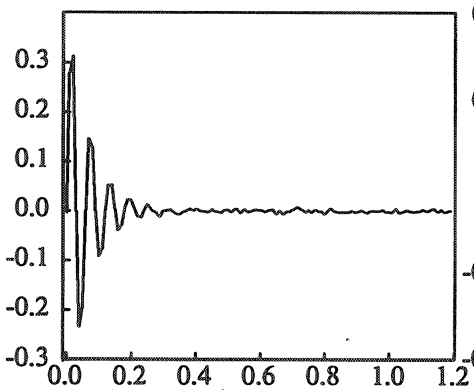

Fig 7.a. Noisy source wavelet

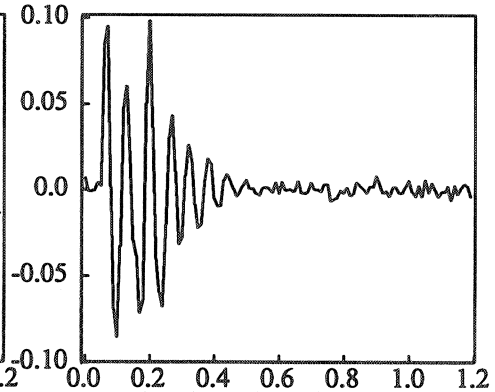

Fig 7.b. Noisy seismogram

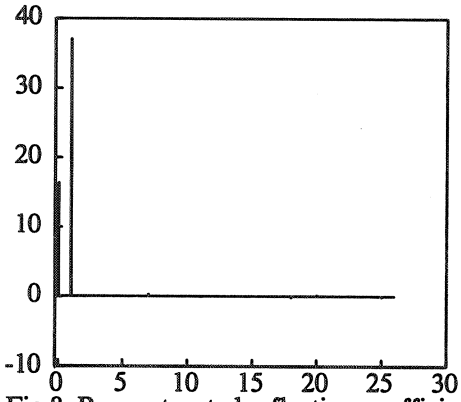

Fig 8. Reconstructed reflection coefficients

EXAMPLE 8.4. Reconstruction through unrestricted ARX parameter estimation.
In the present example and the next example we use the noisy signals of example 8.3, shifted in time. The shift in time is necessary to make sure that an ARX-representation of type (9) makes sense. The shifted noisy signals are depicted in figure 9. In the present example we illustrate the method of reconstructing the reflection coeffcients by the estimation of the parameters of an ARX-representation where no restrictions put on the parameters. Therefore, the function $J(\theta)$ has to be minimized. This is done by means of a quasi Newton minimization method, in which the value of the function and its gradient are computed using algorithm 7.1. The reconstructed reflection coefficients are depicted in figure 10. Figure 10 illustrates the fact that, although $J(\theta)$ is minimized accurately, the reflection coefficients computed by (7) may not be physically realistic.

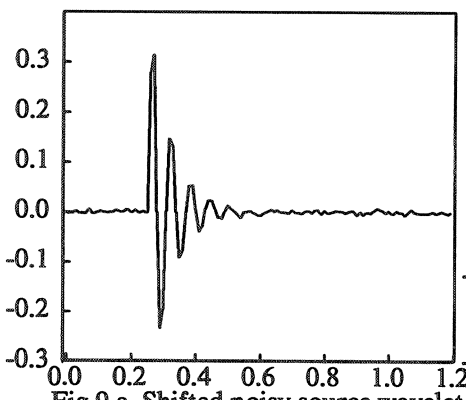

Fig 9.a. Shifted noisy source wavelet

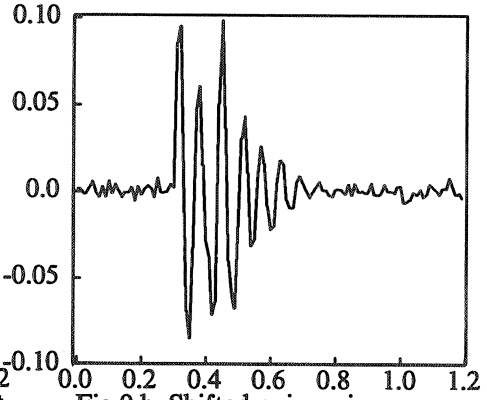

Fig 9.b. Shifted noisy seismogram

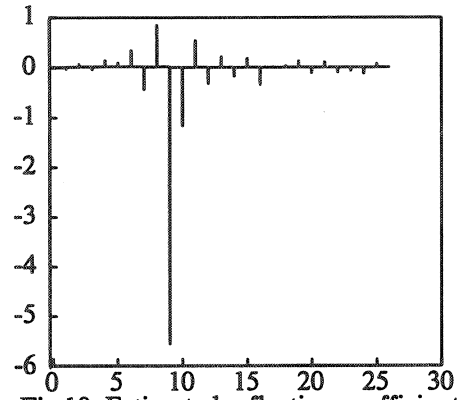

Fig 10. Estimated reflection coefficients

EXAMPLE 8.5. Direct estimation of the reflection coefficients.
In this last example we illustrate the method of the reconstruction of the reflection coefficients by means of the estimation of ARX-parameters that have to correspond to realistic reflection coefficients. Here the function $J(\theta(r(s)))$ has to be minimized. Again, this is done by means of a quasi Newton minimization method, but now the value of the function value and its gradient are computed using algorithm 7.2. The reconstructed reflection coefficients are depicted in figure 11. From figure 11 it is clear that for the reconstruction of the reflection coefficients the method of the estimation of ARX-parameters corresponding to realistic reflection coefficients is to be preferred to the method of the unrestricted estimation of the ARX-parameters followed by (7).

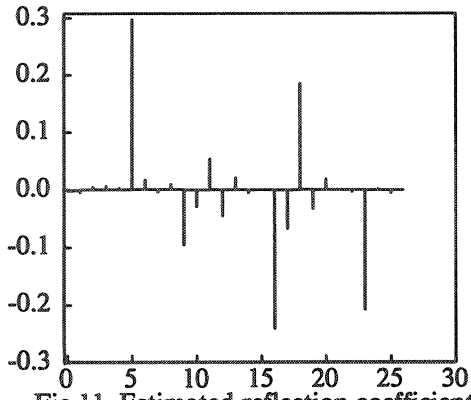

Fig 11. Estimated reflection coefficients

9. CONCLUSIONS AND REMARKS

In this paper we are concerned with the reconstruction of the parameters of a lossless layered medium modelling the interior of the earth. These parameters are also known as the reflection coefficients, and reflect the physical properties of the interior of the earth. To do the reconstruction we use input-output data in the form of the source wavelet and the seismogram, and we make a distinction between the case that the input-output data is noise free, and the case that the input-output data is not noise free.

In the first case we derive a so-called layer stripping method for the reconstruction of the reflection coefficients. The method is just one of the many layer stripping methods that are known, and is illustrated by means of an example. Also the poor performance of the method in the case of noisy input-output data is illustrated by means of an example, indicating that in realistic cases layer stripping

methodes fail to perform properly.

In the second case we develop two new methods which are based on the estimation of the parameters of a so-called ARX-representation. The distinction between the two methods is that in one method the parameters to be estimated are competely free, while in the other method the parameters have to correspond with reflection coefficients that are physically realistic. Both methods are illustrated by means of an example. As a result of the numerical experiments described in section 8, it follows that the estimation of ARX-parameters corresponding to realistic reflection coeficients is to be preferred to the unrestricted estimation of ARX-parameters. A shortcoming of both methods is that they introduce new 'nonzero' reflection coefficients. In many practical situations only a small number of reflection coefficients is actually nonzero. Therefore, it may be worthwhile to try to modify the methods in a such way that the number of nonzero reflection coefficients is estimated, together with their position and magnitude. Also in practical situations the number of layers is much larger than in the examples. Again in these situations the above modification may turn out to be useful. Therefore, in a future investigation it may be interesting to explore how the methods presented in the second part of this paper perform in situations of a small number of nonzero reflection coefficients or of a large number of layers.

REFERENCES

1. M. AOKI and P.C. YUE (1970). On a priori error estimates of some identification methods, *IEEE Trans Automat. Contr.*, AC-15, 541-548.
2. A.M. BRUCKSTEIN and T. KAILATH (1987). Inverse scattering for discrete transmission-line models, *SIAM Review*, 29, 359-389.
3. A.M. BRUCKSTEIN, I. KOLTRACHT, and T. KAILATH (1986). Inverse scattering with noisy data, *SIAM J. Sci. Stat. Comput.*, 7, 1331-1349.
4. K.P. BUBE and R. BURRIDGE (1983). The one-dimensional inverse problem of reflection seismology, *SIAM Review*, 25, 497-559.
5. J.F. CLAERBOUT (1976). *Fundamentals of Geophysical Data Processing*, McGraw-Hill, New York.
6. G. CYBENKO (1980). The numerical stability of the Levinson-Durbin algorithm for Toeplitz systems of equations, *SIAM J. Sci. Stat. Comput.*, 1, 303-319.
7. G.H. GOLUB and C.F. VAN LOAN (1983). *Matrix Computations*, The John Hopkins University Press, Baltimore.
8. F. HABIBI-ASHRAFI and J.M. MENDEL (1982). Estimation of parameters on lossless layered media systems, *IEEE Trans. Automat. Contr.*, AC-27, 31-49.
9. J.M. MENDEL and J. GOUTSIAS (1986). One-dimensional normal-incidence inversion : a solution procedure for band-limited and noisy data, *Proc. IEEE*, 74, 401-414.
10. B. URSIN and K.A. BERTEUSSEN (1986). Comparison of some inverse methods for wave propagation in layered media, *Proc. IEEE*, 74, 389-400.
11. J.M. TEN VREGELAAR (1988). Parameter estimation from noisy observations of inputs and outputs, in *Proceedings of the Third Internatonal Workshop on Statistical Modeling*, 253-279, Vienna.