DEDUCING SUBSURFACE PROPERTY GRADIENTS FROM SURFACE WAVE DISPERSION DATA

John Richardson
Science Center, Rockwell International
Thousand Oaks, California

My talk deals with the same subject as that of the previous speaker, Dr. Szabo, except I am not going to assert that it involves residual stresses necessarily. I consider some kind of a perturbation of the elastic properties and/or density due to some kind of surface treatment. As in all problems of this kind I assume, at least for the time being, some kind of an ansatz about mutual proportionality between the various kinds of perturbed physical properties as they vary with depth.

There are, perhaps, four ways of classifying approaches to this kind of problem. One is a parametric approach where one assumes that the candidate profile is defined by a finite set of parameters which are adjusted to give the best fit to the experimental data. Another is a nonparametric approach in which one does not have a finite set of parameters but an essentially infinite set of parameters. Another dimension of classification is whether one is using a probabilistic or a nonprobabilistic approach. In the nonparametric case one is forced to use a probabilistic approach. Here one treats every conceivable profile as being present in a statistical ensemble but with probability weightings reflecting one's a priori knowledge of what is more or less reasonable.

Dense Data Case

I'm going to start off by considering the dense data case in which the dispersion data is assumed to be given everywhere at all wave lengths - or at least sufficiently dense on the wave length axis that interpolation between the points is not serious. I'm going to first proceed to solve the problem as though the continuous input data were exact and then we will see what kinds of problems are encountered.

I will start off with the same integral equation that Dr. Szabo used, but written in slightly different notation, namely

\[ g(k) = k \int_0^\infty dz \ K(kz) \ f(z) \]  

where \( g(k) \) is the relative change of the Rayleigh velocity at wave number \( k \) due to the subsurface structure described by the profile \( f(z) \) giving a scalar measure of the perturbed material properties as a function of the depth \( z \). The kernel function \( K(kz) \) derived by Tittman and Thompson is a relatively complex function of the unperturbed material properties. A detailed discussion of the kernel and its derivation may be found in Appendix A of a recent paper by myself.
Now, let us "solve" this integral equation in a formal sense. Here, we are not only leaving the real physical world, but as you will see, the real mathematical world also. So, it is a departure of higher order, perhaps appropriate for the last day of a conference.

We observe that the problem has scale invariance, that is, if we change the depth of the profile by a constant scale factor, and if we change the wave lengths by the same scale factor (or equivalently the wave numbers by the reciprocal scale factor) then we obtain another integral equation that is exactly the same form as the one we had before except that it connects the transformed quantities (i.e., transformed dispersion curve and the transformed profile). So in this sense the problem is scale invariant. Now, by considering logarithmic scales rather than the original ones, i.e., we let

\[ z = a e^\beta \]  
\[ k = a^{-1} e^{-\alpha} \]  

we find that the previous scale invariance is converted into translation invariance in the new coordinate system.

This, of course, means in this new coordinate system we obtain the convolution problem:

\[ G (\alpha) = \int \limits_{-\infty}^{\infty} d \beta \ L (\alpha-\beta) \ F (\beta) \]  

where

\[ G (\alpha) = g (a^{-1} e^{-\alpha}) \]  
\[ F (\beta) = f (a e^\beta) \]  
\[ L (x) = e^{-x} K (e^{-x}) \]  

As every electrical engineer knows, this problem can be solved directly by using a Fourier transform. This is not to be confused with the Laplace transform that Dr. Szabo used. He used that in the original z space or k space and so forth. Taking the Fourier transform of this convolution problem, we obtain in the usual way the result:

\[ \mathcal{G} (\xi) = \mathcal{L} (\xi) \mathcal{F} (\xi) \]  

where

\[ \left( \frac{\mathcal{G} (\xi)}{\mathcal{F} (\xi)} \right) = \int \limits_{-\infty}^{\infty} dx \ e^{-i\xi x} \left( \frac{G(x)}{L(x)} \right) \]  

The frequency variable $\xi$ has the same relationship to $a$ or $\beta$ as ordinary frequency does to time in a time-dependent problem. Now we have an algebraic problem whose solution is

$$F(\xi) = \mathcal{L}(\xi)^{-1} \mathcal{G}(\xi)$$

(10)

We then obtain the profile

$$F(\beta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\xi \, e^{i\xi\beta} \mathcal{L}(\xi)^{-1} \mathcal{G}(\xi)$$

(11)

in $\beta$-space or

$$f(z) = F(\log z - \log a)$$

(12)

in the original $z$-space.

This is ill-posed because the integrand of (11) has a factor $\mathcal{L}(\xi)^{-1} \rightarrow \infty$ as $\xi \rightarrow \pm \infty$. Thus, we have a vanishing denominator problem at the end points of the range of integration. Unless the vanishing denominator is compensated by a suitably vanishing function in the numerator, the integral diverges. Then, of course, we get into the question of how should the function in the numerator be made to vanish? If it were an analytic function which vanished sufficiently rapidly as $\xi \rightarrow \pm \infty$, we could obtain a finite integral. Starting with real data, suitable interpolation, extrapolation and smoothing would be required to produce such a function. Unfortunately, the value of the integral would depend crucially on the particular parameter values characterizing these operations. Thus, the problem in the present form is ill-posed.

This immediately indicates an estimation theory approach which is capable of handling problems of both noisy and incomplete data. In the dense-data case now under discussion, the problem is a matter of fighting noise, not incompleteness of data and we thus assume a model which involves the same integral equation we had before except that it has been recast in a probabilistic context. We thus write

$$G(a) = \int_{-\infty}^{\infty} d\beta \, L(a-\beta) F(\beta) + \nu(a)$$

(13)

where $\nu(a)$ is the measurement noise and where $F(\beta)$ is now regarded as a random process.

We thus have a statistical ensemble of conceivable functions $F(\beta)$ representing all the moderately likely (and unlikely, too, for that matter) profiles and we have another ensemble representing the noise. We assume that $F$ and $\nu$ are Gaussian random processes with the following means and covariances:

$$E F(\beta) = F^0(\beta)$$

(14)

$$E \Delta F(\beta) \Delta F(\beta') = C_F (\beta-\beta')$$

(15)

$$E \nu(a) = 0$$

(16)
\[ E v(a) \nu(a') = C_{v}(a-a') \]  
(17)

\[ E \Delta F(\beta) \nu(\alpha) = 0 \]  
(18)

where \( E \) is the averaging operator and \( \Delta \) is defined by

\[ \Delta(\cdot) = (\cdot) - E (\cdot). \]  
(19)

This combined ensemble, which represents our a priori statistical knowledge of the situation, must be hedged in by the data obtained from a set of experiments. This is called conditioning. We thus throw out the members of the statistical ensemble which are inconsistent with the measurements and, in this remaining reduced ensemble, we look for the most probable profile. This is also equivalent to using an optimal estimator which is a function of the measured data \( \tilde{G}(\alpha) \), namely

\[ \hat{F}(\beta; \tilde{G}(\alpha)). \]  
(20)

This gives the best estimate of \( F(\beta) \) when the actual measured value of \( G(\alpha) \) is \( \tilde{G}(\alpha) \). The optimal estimator is given by the functional form that minimizes the mean square criterion

\[ \varepsilon = \frac{1}{2} E \left( \hat{F}(\beta; G(\alpha)) - F(\beta) \right)^2 \]  
(21)

It is to be noted that in the optimization we have used the model process \( G(\alpha) \) is defined by (13) and not the actually measured dispersion curve \( G(\alpha) \). Thus, the estimator is optimized in accordance with what works best for the model in a mean square sense.

Since the measurement process is linear in both the measurement error \( \nu(\alpha) \) and the profile \( F(\beta) \) and since these functions are Gaussian, the optimal estimator is linear in \( G(\alpha) \). We thus assume a candidate estimator of the form

\[ \hat{F}(\beta; G(\alpha)) = b(\beta) \]

\[ + \int_{-\infty}^{\infty} d\alpha B(\beta-\alpha) \Delta G(\alpha) \]  
(22)

The optimal linear estimator is then

\[ \hat{F}(\beta; G(\alpha)) = E F(\alpha) \]

\[ + \frac{1}{2\pi} \int_{-\infty}^{\infty} d\xi e^{i\xi \beta} \overline{\Lambda}(\xi)^{-1} \Delta G(\xi), \]  
(23)

where \( \overline{\Lambda}(\xi) \) is defined by

\[ \overline{\Lambda}(\xi) = \left[ 1 + |R(\xi)||U(\xi)|^2 \right]^{-1}, \]  
(24)
and in turn $\mathbb{R}(\xi)$ is given by

$$\mathbb{R}(\xi) = \overline{C}_v(\xi) \overline{C}_f(\xi)^{-1}$$

(25)

We will use $\mathbb{I}(\alpha) \mathbb{R}(\alpha)^{-\frac{1}{2}}$ as a measure of the signal-to-noise ratio which we will frequently denote by the expression $S/N$.

The optimal estimate in (23) is given by the a priori average profile $EF(\beta)$ plus the integral in $\xi$-space representing a correction due to the observational conditioning. This is to be compared with the formal solution we had before.

We can rewrite it in the form

$$F(\beta) = EF(\alpha) + \frac{1}{2\pi} \int_{-\infty}^{\infty} d\xi \, e^{i\xi \beta} \, \mathbb{I}(\xi)^{-1} \triangle G(\xi)$$

(26)

The difference between (23) and (26) is that in (23) a convergence factor $\overline{\Lambda}$ has appeared. This convergence factor is not some ad-hoc thing evoked just for the sake of finiteness, but comes directly out of the above statistical model by the processes of estimation theory. This convergence factor completely compensates for the vanishing denominator giving a well-posed solution.

We can write the estimator in the alternative form

$$\hat{F}(\beta; G(\alpha)) = E \, F(\beta) + \int_{-\infty}^{\infty} d\alpha \, Q(\beta-\alpha) \triangle G(\alpha),$$

(27)

where

$$Q(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\xi \, e^{i\xi x} \, \overline{\Lambda}(\xi) \, \mathbb{I}(\xi)^{-1}.$$  

(28)

Here the correction to the a priori average is an integral in $\alpha$-space involving a kernel $Q(\beta-\alpha)$ which I will call the quasi-inverse of $L(\alpha-\beta)$.

The term "quasi-inverse" comes from the following relation

$$\int_{-\infty}^{\infty} d\alpha \, Q(\beta-\alpha) \, L(\alpha-\beta') = \Lambda(\beta-\beta')$$

(29)

where

$$\Lambda(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\xi \, e^{i\xi x} \, \overline{\Lambda}(\xi).$$

(30)

The term $\Lambda(\beta-\beta')$ on the r.h. side of (29) is an approximation to a $\delta$-function. Since $Q(x)$ is defined in terms of $\overline{\Lambda}(\xi)$ by (30), we see that $\Lambda(\xi)$ does indeed become a $\delta$-function as $\overline{\Lambda}(\xi) \rightarrow 1$, i.e., in the limit of infinite signal-to-noise ratio.
It is interesting to see how \( Q(\beta - \alpha) \) behaves as a function of the signal-to-noise ratio. At low values of this ratio \( Q \) is fairly well behaved as illustrated in Fig. 1a. This function is convolved with the dispersion curve in order to get the estimate of the profile in the case where the \textit{a priori} average of the profile vanishes. I want to emphasize that one unit on the horizontal scale corresponds to a factor of \( e \) in depth, i.e., it is a logarithmic function of depth. For a high signal-to-noise ratio, we get a function that is highly oscillatory as seen in Fig. 1b, and, of course, highly vulnerable to noise.

As you go on to higher and higher signal-to-noise ratios, \( Q \) approaches mathematical nonexistence. This function becomes more and more highly oscillatory and this is the typical kind of situation one encounters in doing deconvolution.

We now consider estimates based upon theoretical experiments. Figure 2 shows a theoretical experimenter who has the right answer which he puts into the computer, which generates theoretically measured data which comes sifting down through a hole in the floor. Below, a diligent fellow called an estimator infers from the numbers falling through the hole what the profile should be according to estimation theory. I want to emphasize this in order to keep straight what is going on in the production of theoretical test data and what is going on in the estimation process. It is happening on two different floors.

Let us assume that the right (i.e., assumed) profile is \( f(z) \) to which there corresponds a profile \( \tilde{F}(\beta) \) in \( \beta \)-space i.e., \( \tilde{F}(\beta) = \tilde{f}(ae^\beta) \). If the test data is assumed to be noiseless, then the synthetic test data is given by

\[
\tilde{G}(\alpha) = \int_{-\infty}^{\infty} d\beta \ L(\alpha - \beta) \tilde{F}(\beta)
\]

The estimated profile in \( \xi \)-space is with the use of (27) found to be

\[
\tilde{F}(\xi; \{\tilde{G}(\alpha)\}) = (1 - \tilde{\Lambda}(\xi)) \tilde{E}(\xi)
+ \tilde{\Lambda}(\xi) \tilde{F}(\xi)
\]

where \( \tilde{E}(\xi) \) and \( \tilde{F}(\xi) \) are the Fourier transforms of \( E(\beta) \) and \( F(\beta) \) using obvious extensions of (9).

If the right answer happens to be a \( \delta \)-function, the estimator comes up with the kind of a profile shown in Fig. 3a. The result is hardly a \( \delta \)-function, but it is his best attempt to get one for a signal-to-noise ratio of 3.16. At a much higher signal-to-noise ratio one obtains a much better approximation to a \( \delta \)-function, at least as far as the central lobe is concerned, but with a somewhat busier side lobe structure. In Fig. 3b we give an example for a signal-to-noise ratio of 100.

The widths of the approximate \( \delta \)-functions illustrated in the last two figures define the resolution of the estimation process as a function of signal-to-noise ratio. To be more precise, the resolution here is really the fractional
Fig. 1. The quasi-inverse $Q(X)$; (a) $S/N = 3.16$ and (b) $S/N = 100$.  

775
Fig. 2. Preparation of synthetic test data.
Fig. 3. The quasi-δ-function $\Lambda(X)$; (a) $S/N = 3.16$ and (b) $S/N = 100$. 
resolution, i.e., the width of the approximate $\delta$-function as a function of depth $z$ (instead of the logarithmic variable $\beta$) divided by the average depth. In other words, an approximate $\delta$-function extending between 1.0 and 1.1 $\mu$ and another between 1.0 and 1.1 cm represent the same fractional resolutions. In Fig. 4 we show the variation of the fractional resolution $\lambda$ as a function of the $S/N$. Clearly, the attainment of a very fine fractional resolution represents a stiff uphill battle in terms of $S/N$. In particular, as $S/N$ goes from 3.16 to 100, one obtains a reduction of $\lambda$ of only a little better than a factor of 2.

**Sparse-Data Case**

We turn now to a discussion of the sparse-data case, the one of practical importance. Here, we employ a model that is essentially the same as before except that we consider a discrete set of dispersion data instead of a continuous curve as was done in the dense-data case. The appropriate probabilistic version of (1) is now

$$g_n = \int_{-\infty}^{\infty} dz J_n(z) f(z) + \nu_n, \quad n = 1, \ldots, N,$$

where $g_n$ and $J_n(z)$ are defined by

$$g_n = g(k_n^0)$$

and

$$J_n(z) = k_n K(k_n z)$$

The quantity $k_n$ is the value of the wave number $k$ at which is made the $n$th measurement of the relative change $g(k_n^0)$ of the Rayleigh velocity.

As before, we assume (a priori) that the possible profile $f(z)$ is a Gaussian random process and that the measurement errors are Gaussian random variables. We now assume the properties

$$E f(z) = f^0(z),$$

$$E \nu_n = 0,$$

$$E \Delta f(z) \Delta f(z') = C_f^0 \delta(z-z'),$$

$$E \nu_n \nu_n' = \delta_{nn'} \sigma^2,$$

$$E \Delta f(z) \Delta \nu_n = 0.$$

The statistical properties of $f(z)$ assumed here differ from those assumed in the previous case in that $f(z)$ is stationary with respect to $z$, not $\log z$. We assume for convenience that $f(z)$ is a white random process corresponding to the absence of an a priori smoothness bias.
Fig. 4. Variation of the fractional resolution $\lambda$ with the signal-to-noise ratio $S/N$. 

$$\lambda = \frac{\int dx \Lambda(x)}{\Lambda(x)_{\text{max}}} = \frac{\Lambda(0)}{\Lambda(0)}$$
The best estimator in a least mean square sense is given by

$$\hat{f}(z; \{\tilde{g}_n\}) = Ef(z)$$

$$+ C_f^0 \sum_{n,n'} J_n(z) C^{-1}_{g,n,n'} (\tilde{g}_n - Eg_n)$$

(39)

where $C^{-1}_{g,n,n'}$ is the matrix inverse of the covariance matrix of the $g_n$, namely,

$$C_{g,n,n'} = E g_n g_n'$$

$$= M_{n,n'} + \sigma^2 \delta_{n,n'},$$

(40)

where

$$M_{n,n'} = C_f^0 \int_0^{\infty} dz J_n(z) J_n(z).$$

(41)

The a priori average of $g_n$ is given by

$$E g_n = \int_0^{\infty} dz J_n(z) E f(z)$$

(42)

It is worthy of note that the quantity $f(z; \{\tilde{g}_n\}) - E f(z)$ must always lie in the subspace spanned by the $J_n(z)$. Thus, once the experimentalist has specified a set of wave numbers $n_k$, we have a universal subspace in which all estimates (after subtraction of the a priori average $E f(z)$) must lie. Outside of this subspace the observational conditioning provides no correction of $E f(z)$ to give the estimate. In other words, outside of this subspace the estimate is completely model dominated while inside it is data dominated to a degree dependent upon the variances of the observational errors.

It is of interest to consider a special basis set for the above subspace. The first step in the derivation of this set is the solution of the following eigenvalue-eigenvector problem

$$\sum_{n=1}^{N} M_{n,n'} U_{n',\alpha} = U_{n,\gamma} \alpha, \quad n,\alpha = 1, \ldots, N$$

(43)

subject to the orthonormality constraints,

$$\sum_{n=1}^{N} U_{n,\alpha} U_{n',\alpha} = \delta_{\alpha \alpha'},$$

(44)
We then define the new basis functions by the relation

\[ \psi_\alpha(z) = (C_\alpha^t)^{1/2} \gamma_\alpha^{1/2} \sum_{n=1}^{N} U_{n\alpha} J_n(z) \]  \hspace{1cm} (45)

It is easy to show that the \( \psi_\alpha(z) \) satisfy the orthonormality conditions

\[ \int_{a}^{b} \psi_\alpha(z) \psi_\beta(z) \, dz = \delta_{\alpha\beta}, \]  \hspace{1cm} (46)

In Fig. 5 we show examples of the \( \psi_\alpha \) for the case of 4 measurements, i.e., \( N=4 \). The values of wavelength and wavenumber are given in Table I.

<table>
<thead>
<tr>
<th>n</th>
<th>Wavelength</th>
<th>( k_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.050 cm</td>
<td>125.6 cm(^{-1})</td>
</tr>
<tr>
<td>2</td>
<td>0.075</td>
<td>83.78</td>
</tr>
<tr>
<td>3</td>
<td>0.295</td>
<td>21.30</td>
</tr>
<tr>
<td>4</td>
<td>0.595</td>
<td>10.56</td>
</tr>
</tbody>
</table>

The values of the eigenvalues \( \gamma_\alpha \) of the matrix \( M_{nn} \) are presented in Table II.

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( \gamma_\alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.401</td>
</tr>
<tr>
<td>2</td>
<td>0.987</td>
</tr>
<tr>
<td>3</td>
<td>0.327</td>
</tr>
<tr>
<td>4</td>
<td>0.109</td>
</tr>
</tbody>
</table>

The unperturbed material properties used in computing the \( J(z) \) are those of 1043 steel. In relating \( f(z) \) to perturbed material properties we have assumed that the surface layer hardening process leaves the compressibility and the density unaltered.

It is to be noted that the functions \( \psi_\alpha(z) \) increase in complexity as \( \alpha \) increases (i.e., as \( \gamma_\alpha \) decreases). Examination of the curves in Fig. 5 reveals that the number of stationary points (for finite values of \( \log z \)) is equal to \( \alpha + 1 \).

Let us extend the finite set \( \psi_\alpha(z), \alpha = 1, \ldots, N \), into a complete set by adding other normalized functions that are orthogonal to the members of the original set and to each other. A possible profile can be expressed in the form
Fig. 5. Basis functions \( \psi_{\alpha}(z) \).
\[ F(z) = \sum_{\alpha=1}^{N} f_{\alpha} \psi_{\alpha}(z) + \sum_{\alpha=N+1}^{\infty} f_{\alpha} \psi_{\alpha}(z) \]  

(47)

where the coefficients are obviously given by

\[ f_{\alpha} = \int_{0}^{\infty} dz \psi_{\alpha}(z) f(z). \]  

(48)

These coefficients are random variables that can be shown to be statistically independent both a priori and a posteriori. The former assertion can be proved by substituting (47) into (36)—thereby obtaining an expression from which we readily infer the result

\[ E \Delta f_{\alpha} \Delta f_{\alpha'} = C^0_f \delta_{\alpha \alpha'} . \]  

(49)

The proof of the second assertion is more difficult and hence we will give the results with the omission of the intervening steps. We obtain the a posteriori (i.e., observationally conditioned) covariance matrix in the form

\[ E(\Delta_g f_{\alpha} \Delta_g f_{\alpha'}) \mid \{g_n\} = (1 - \Lambda_{\alpha}) C^0_f \delta_{\alpha \alpha'} , \]  

(50)

where

\[ \Lambda_{\alpha} = \frac{\gamma_{\alpha}}{\gamma_{\alpha} + \sigma} , \quad \alpha = 1, \ldots, N \]

\[ = 0, \quad \alpha > N. \]  

(51)

The symbol \( \Delta_g \), denoting the deviation from a posteriori average is defined by

\[ \Delta_g(\cdot) = (\cdot) - E(\cdot \mid \{g_n\}) \]  

(52)

The quantity \( \Lambda_{\alpha} \) measures the degree of data dominance with respect to the basis function \( \psi_{\alpha}(z) \). One notes that \( \Lambda_1 \) is the largest of the \( \Lambda_{\alpha} \) and from \( \Lambda_2 \) to \( \Lambda_N \) the magnitudes are nondecreasing. If the variance of the measurement vanishes, i.e., \( \sigma = 0 \), the quantities \( \Lambda_1, \ldots, \Lambda_N \) are all equal to unity. When \( \alpha > N \), the \( \Lambda_{\alpha} \) all vanish corresponding to complete model dominance.

The estimator can be expressed in an interesting way in terms of the basis functions \( \psi_{\alpha}(z) \). We can write

\[ \hat{f}(z; \{g_n\}) = \sum_{\alpha=1}^{\infty} \hat{f}_{\alpha} \psi_{\alpha}(z) \]  

(53)
where the estimated coefficients are given by

\[ \hat{f}_\alpha = (1 - \Lambda \alpha) E f + \Lambda \alpha \tilde{g}_\alpha. \]  

(54)

The quantities \( \tilde{g}_\alpha \) are linear combinations of the \( g_n \) defined by

\[ \tilde{g}_\alpha = (\gamma^{a}_{f}) \sum_{n=1}^{N} U_{n\alpha} g_n. \]  

(55)

Let us consider the use of noiseless (i.e., no measurement errors) theoretical test data based upon an assumed profile \( f(z) \). We obtain the test data in the form

\[ \tilde{g}_\alpha = f_\alpha \]  

(56)

where, of course,

\[ \tilde{f}_\alpha = \int_{0}^{\infty} dz \, \psi_\alpha(z) \tilde{f}(z). \]  

(57)

The estimated \( f_\alpha \) is thus according to (54) given by

\[ \hat{f}_\alpha = (1 - \Lambda \alpha) E f + \Lambda \alpha \tilde{f}_\alpha \]  

(58)

which is closely analogous to (30b) in the previous dense-data case.

We have considered an assumed profile of the form

\[ \tilde{f}(z) = B \exp \left( - \frac{z}{b} \right). \]  

(59)

In Fig. 6 we present the comparison of the estimated and assumed profiles, \( \hat{f}(z) \) and \( \tilde{f}(z) \), respectively, for the case of \( B = 1 \) and \( b = 0.1 \) cm. The theoretical test data \( \tilde{g}_\alpha \) were calculated with the \( k_n \) values listed in Table I and the unperturbed material properties corresponding to 1043 steel. In the estimator we have assumed the values \( \sigma = 0 \) and 0.3 for the standard deviation of the measurement error. The agreement between the estimated and assumed curves is about approximately what one should expect for 4 data points and no a priori smoothness bias.

Of particular interest is the case where the assumed profile is a \( \delta \)-function, i.e.,

\[ \tilde{f}_\alpha(z) = \delta(z-z') \]  

(60)

where upon using (57) we obtain

\[ \tilde{f}_\alpha = \psi_\alpha(z'). \]  

(61)
Fig. 6. Comparison of estimated profiles $\hat{f}$ with assumed profiles $f$.
(a) $\sigma = 0$ and (b) $\sigma = 0.3$. 
Then (58) reduces to

\[ \hat{f}_\alpha = (1 - \Lambda_\alpha) E f_\alpha + \Lambda_\alpha \psi_\alpha(z'). \]  

(62)

In the original z-space, the above results go into

\[ \hat{f}(z; \{ \tilde{g}_n \}) = \int_0^\infty dz'' \left[ \delta(z-z'') - \Lambda(z,z'') \right] E f(z'') + \Lambda(z,z') \]  

(63)

where

\[ \Lambda(z,z') = \sum_{\alpha=1}^R \psi_\alpha(z) \Lambda_\alpha \psi_\alpha(z'). \]  

(64)

In the case where \( Ef(z)=0 \), the estimator (63) reduces simply to

\[ \hat{f}(z; \{ \tilde{g}_n \}) = \Lambda(z,z') \]  

(65)

In Fig. 7 we show the comparison of the estimated profile (65) with the assumed profile (60) for several values of \( z' \). In the first two plots, the failure of \( \hat{f} \) to approach small values as \( z \to \infty \) is not serious since below \( z = 0.005 \) cm, say, the area under the curve is negligible compared with the area under the main peak.

We turn now to estimation based upon real data. The dispersion data provided by Tittmann (Table III) have

\[ \begin{array}{cccc}
 n & \text{Wavelength} & k_n & v_R + \delta v_R (k_n) \\
 \hline
 1 & 0.050 \text{ cm} & 125.6 \text{ cm}^{-1} & 2.934 \times 10^5 \text{ cm sec}^{-1} \\
 2 & 0.075 & 83.78 & 2.9344 \\
 3 & 0.295 & 21.30 & 2.946 \\
 4 & 0.595 & 10.56 & 2.969 \\
 \end{array} \]  

been introduced into (39) to yield the solid curve in Fig. 8. The open circles represent the results of independent destruction measurements. In the estimator we have assumed \( Ef(z)=0, \sigma=0 \), no a priori smoothing bias and a set of unperturbed material properties corresponding to 1043 steel.
Fig. 7. Comparison of estimated profile $\hat{f}$ with assumed $\delta$ function profiles $\tilde{f}$; (a) $z' = 0.0067$; (b) $z' = 0.0159$; (c) $z' = 0.0411$; and (d) $z' = 0.0167$. 
Fig. 8 Estimated profile vs. experiment.
Summary

We have presented the non parametric treatment of the Rayleigh inverse problem using an estimation theory approach. We have considered both the dense-data and sparse-data cases. In one the estimator was beset by noise in the data and in the second case the estimator was beset by incompleteness of data. I think these results give you a feeling for the kind of problems that you are fighting against in trying to obtain good estimates with the kind of data that occurs in the real world.

References


5. B. R. Tittmann, private communication.
DISCUSSION

DR. WALKER (AFOSR): I think we have time for just one question.

DR. DON THOMPSON (Rockwell International, Science Center): John, I think you might comment on how real world that data is. It came from a truck axle.

DR. RICHARDSON: That one?

DR. DON THOMPSON: Yes.

DR. RICHARDSON: That is as real as you can get. I also want to comment that the estimator that I had for the sparse data case looks complicated, but the actual calculations in the field (which could be performed by a high school graduate) are very simple. The on-line part is elementary and the off-line part is relatively complicated. Well, thank you very much.