

on the reflector, the *singular function of the reflecting surface*, as depicted in Figure 5. The scaling factor of the singular function is the *normal incidence geometrical optics reflection coefficient*. This product of singular function with reflection coefficient is what we call the *reflectivity function* for this theory.

In 1983, I was fortunate enough to spend two months at Amoco's Tulsa Research Center, then one of the premier exploration research centers in the world. My last lecture there was entitled, "How to Invert Anything." In that talk, I proposed the following procedure for determining the weights.

1. Apply the proposed inversion formula to Kirchhoff-approximate data for the upward scattered field.
2. Carry out all integrals but the  $\omega$ -integral (implicit in the Fourier filtering of the data) asymptotically, by the method of stationary phase.
3. Choose the frequency domain filter so that the only frequency dependence left in this last integral is the source signature. (This assures that the frequency domain integral is a bandlimited delta function.)
4. If possible, choose  $B$  so that the resulting amplitude is proportional to a reflection coefficient.

I will explain below how the peak of that delta function turns out to always occur right on the reflector. Accepting that for a moment, as soon as the frequency domain filter is determined, we are assured that the output will be a *reflectivity map* of the surface. All one needs to do, thereafter, is choose any weighting function that produces reasonable balancing of intensity for shallow and deep reflectors. That is Kirchhoff migration for a general, variable background medium for a prescribed source/receiver configuration.

In fact, this procedure was used to produce a common offset, constant background, inversion in Sullivan and Cohen [1987]. However, this approach to deriving inversion formulas was both tedious and non-insightful. Each new source/receiver configuration and background required analysis of the determinant of a matrix of second derivatives of the phase function.

Luckily, while we were struggling with this approach, in a seminal paper, Gregory Beylkin [1985] proposed a general method that obviated the need for treating each source/receiver configuration separately. The mathematical tools behind this general method include *pseudo-differential operator theory* and its natural partner, *Fourier integral operator theory*. Beylkin used this method to propose an inversion for velocity perturbation. However, in processing for this perturbation, we would be relying on bandlimited high frequency data to produce a reliable image of steps, the discontinuities of the medium parameters across reflectors. High frequency data does not do this very well. On the other hand, the derivatives of those steps are delta functions. High frequency data does a far better job of processing for those!

From all of the struggling we had done, we knew, immediately how to transform Beylkin's inversion for velocity perturbation into our inversion for a reflectivity function and we proceeded to do so in a series of papers: Cohen and Hagin [1985]; Cohen, Hagin and Bleistein

[1986]; and Bleistein, Cohen and Hagin [1987]. Although these results were couched in terms of the Born approximation, suggesting linearized accuracy in the medium perturbations, a subsequent proof showed that the output was actually linear in the reflection coefficient [Bleistein, 1987a,b] and was just the reflectivity function defined above. Hence, we feel justified in using the name *Kirchhoff inversion*, both because of the structure of the operator and for the nature of the output. Furthermore, we were able to confirm that our one-at-a-time approach would have led to exactly the same inversion formulas as did Beylkin’s global approach.

It is also important to note that a part of the weighting in Beylkin’s inversion was the sort of multiplier that could never have been predicted by a migration approach to this problem. Although this factor was present in our approach, as well, it took this pseudo-differential operator approach to explain its role. From that theory, one can deduce that the integration over sources and receivers and frequency is really an approximate inverse Fourier transform. The Fourier variables are, in some sense, dual wave-vector variables to the spatial variables of the forward model obtained either through the Kirchhoff approximation or the Born approximation. The wave-vector is defined locally for each output point,  $\mathbf{x}$ , through the relationship,

$$\mathbf{k} = \omega \nabla_{\mathbf{x}} \tau(\mathbf{x}, \bar{\mathbf{x}}).$$

That is,  $\mathbf{k}$  is just the gradient of the traveltimes from source to output point to receiver, multiplied by the frequency. The rule for making such a change of variable of integration from  $\mathbf{k}$  to  $(\omega, \bar{\mathbf{x}})$  is to multiply by the *Jacobian* of that transformation. In this case,

$$dk_1 dk_2 dk_3 = \omega^2 |h(\mathbf{x}, \bar{\mathbf{x}})| d\omega d\bar{x}_1 d\bar{x}_2.$$

This factor,  $h$ , appears in the inversion formulas that can be found in the literature. From the first paper that we wrote about adapting Beylkin’s method, we referred to  $h$  as the Beylkin determinant, in recognition of the significance of this factor for inversion; that name is now in general use.

Of course, the form of  $h$  changes with each source/receiver configuration and the actual function depends, as well, on the background wavespeed and attendant solutions of the eikonal equation that make up the traveltimes from source and receiver to the output point. In particular, this factor is extremely cumbersome in 3D common offset processing.

### **Kirchhoff migration, why does it work?**

I will offer a mathematician’s explanation of why this processing works. That goes back to the test of applying the basic processing formula to model data—Kirchhoff approximate data—and going through the stationary phase calculations that give an analytical description of how the method works on data. Let us concentrate on the 2D case, so that we can relate what we have done back to the original discussion of the Hagedoorn construction. That Kirchhoff approximation for reflection data is an integral over the curve that defines the reflector. We would want to carry out that integration and the integration over  $\bar{\mathbf{x}}$  by the method of stationary phase. When we apply that method to the integral over the reflector,

we find that, for each source/receiver point, the stationary point on the reflector is just the point where specular reflection occurs, that is, where Snell’s law is satisfied. Further, the phase of the model data will be evaluated at the specular traveltime from source, to reflection surface specular point, to receiver. In fact, this is just the traveltime represented by  $t(\bar{x})$  in (1) and the integrand has exactly the structure of (2)! We have now come full circle. Below that last equation, we described how the combination of stationary phase in  $\bar{x}$  and integration over  $\omega$  leads to an integral representation of a bandlimited delta function that peaks on the reflecting curve—what we can now call the singular function of the curve. The rest, how the amplitude comes out “right” is a matter of details that can be found in the references. Here, “right” means that the singular function is scaled by the reflection coefficient,  $R(\mathbf{x}, \theta)$ , with  $\theta$  a distinguished incidence angle. The distinguished angle is the one for which the incidence directions of the rays from the source/receiver pair satisfy Snell’s law at the imaged point. Remarkably, the theory also provides a method for determining that distinguished angle for little additional cost in computation. This detail can also be found in the references.

### 2.5D inversion.

While Kirchhoff migration and inversion were being developed, single line processing was standard. In this case, it is not reasonable to try to develop a 3D reflectivity. The reason is familiar; out-of-plane reflections cannot be distinguished from in-plane reflections, with the plane in question being the vertical  $(x_1, x_3)$ -plane below the acquisition line. Indeed, consider the surface of revolution obtained by spinning the curve of Figure 1 around the acquisition line. A reflector tangent to any point on a circle of revolution of this surface has the same arrival time and trace location of its specular return as the point on that circle with  $x_2 = 0$ ; that is, at the point on the original generating curve.

As a practical matter, then, in acquiring a single line of data for processing, it would be advantageous to place that line in the dominant dip direction of the subsurface, thereby minimizing reflections from points out of the vertical plane below the acquisition line. The mathematical challenge, then, is to devise a formalism that allows one to process such data and obtain a result that acknowledges the 3D propagation characteristics of the wavefield, but seeks only a 2D description of the subsurface. This formulation and the subsequent processing have acquired the label, 2.5D. I originally heard that term used by Gerry Hohlman, referring to E&M data processing, and I adopted it for geophysical data processing.

The basic idea is to consider the following “thought experiment.” Let us idealize the dip direction by assuming that the earth has no variation in the out-of-plane direction, referring to the vertical plane below the acquisition line. See Figure 6. For this configuration, all parallel lines of data, *with the same acquisition geometry as the given line*, are the same. Thus, if we acquire one line of data for (11), we have acquired the data for the same source/receiver configuration on all lines. That is, the data is independent of the out-of-plane variable,  $\bar{x}_2$ , and only the operator kernel depends on this variable. Of course, the background medium, with respect to which we compute traveltimes and amplitudes also have special features: see Bleistein [1986]; Bleistein, Cohen and Hagin [1987]; and Docherty [1987].

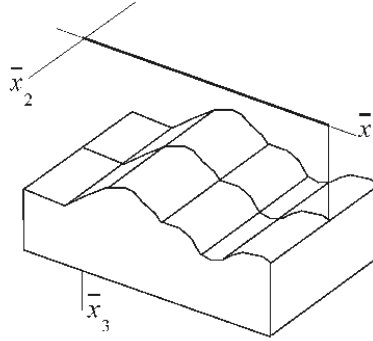


FIG. 6. The black line depicts a source-receiver array located over a 2.5-D model. All specularly reflected ray paths are confined to the plane,  $\bar{x}_2 = 0$ .

These special features allow us to carry out the integration in the  $\bar{x}_2$  direction by the method of stationary phase. Not surprisingly, the stationary value of  $\bar{x}_2$  turns out to be just  $x_2$ , and no dependence on the out-of-plane variables survives the asymptotic analysis. Consequently, the inversion processing reduces to an integration over a single data line. However, the Fourier filter and the spatial weighting for this 2.5D processing is *different* from the filter and weight of 2D processing.

To satisfy the reader's curiosity, I will tell you what those Fourier domain filters turn out to be

$$\begin{array}{ll}
 \text{3D} & i\omega \\
 \text{2.5D} & \sqrt{|\omega|}e^{i\pi/4\text{sign}(\omega)} \\
 \text{2D} & |\omega|.
 \end{array}$$

The 3D filter is equivalent to differentiation in the time domain. The 2.5D filter is equivalent to the Hilbert transform of the half derivative in the time domain through the factorization,

$$\sqrt{|\omega|}e^{i\pi/4\text{sign}(\omega)} = e^{i\pi/2\text{sign}(\omega)} * \sqrt{|\omega|}e^{-i\pi/4\text{sign}(\omega)}$$

with the first factor being equivalent to the Hilbert transform and the second factor, the half derivative. The 2D filter is equivalent to the Hilbert transform of the time derivative because

$$|\omega| = e^{i\pi/2\text{sign}(\omega)} * |\omega|e^{-i\pi/2\text{sign}(\omega)} = e^{i\pi/2\text{sign}(\omega)} * (-i\omega).$$

Again, the first factor yields a Hilbert transform, but now the second factor yields an ordinary first derivative in time. We remark that, in the 2.5D case, while the frequency domain filters commute, the time domain filters do not. Half derivatives are only defined for causal

functions—functions that turn on at a finite time. The Hilbert transform creates a function that is not causal. Hence, if that filter were applied first, the fractional derivative could not then be applied in the time domain. In contrast, in 2D, the full derivative is defined even when the function to which it is applied is not causal, so the order of applying the filters in the time domain is not important.

### **Migration versus inversion: philosophy.**

Migration is based on secondary principles that do not follow directly from the model of the experiments. Let me explain in the context of zero offset data. The most serious of these secondary principles is the following: “The data of the set of zero offset traces can be treated as a wavefield that satisfies the wave equation and, therefore, can be extrapolated back into the earth to zero time as long as we use half the wavespeed of the medium.” *This premise is not true* unless we modify it as follows: “The data of the set of zero offset traces can be treated as a wavefield that *approximately* satisfies the wave equation for constant wavespeed *to leading order at high frequency* and, therefore, can be extrapolated back into the earth to zero time *to this degree of accuracy* as long as we use half the wavespeed of the medium.”

Indeed, each trace is the solution of a *different wave equation*. What differs in each equation is the location of the source; it moves with the observation point so that the two are coincident. There is no *a priori* reason to assume that this ensemble of data from different solutions of the wave equation is, itself, an exact solution of the wave equation. In fact, I never could find a proof in the seismic literature before the attempt proposed in Bleistein and Cohen [1982]. There, we were able to show what I have claimed above. Further, we showed that in this constant wavespeed case, this solution was the response to a source distribution that resided on the reflector. Thus, both *wave equation migration* and the *exploding reflector model* were simultaneously confirmed as leading order asymptotic results by this analysis. Further, for variable background, we showed that there was an additional source term. However, that one was not distributional in nature (as the exploding reflector source is) and, hence, at worst, contributes some low level background field that would not interfere with the identification of reflectors and, at worst, would only degrade the amplitude a bit.

On the other hand, inversion starts from a mathematical model representation of the *ensemble* of traces and proceeds to develop an inversion without ever assuming that the *ensemble* is a solution of the wave equation. High frequency techniques further use linearization as well as high frequency models of the forward wave propagation to produce direct and deterministic formulas for the solution of the inverse problem—but these are leading order asymptotic solutions for which the underlying approximations are fairly well understood. In either case—the migration approach or the inversion approach—the result is a high frequency approximation of the reflector map in the subsurface. The added advantage of the inversion approach is that it predicts a model-consistent relationship between the amplitude of the output and the normal incidence reflection coefficient.

Further, the inversion approach lends itself more easily to the determination of the reflector map for problems with separated source and receiver and nonconstant wavespeed. In

those problems, the amplitude is related in a model-consistent way to the angularly dependent reflection coefficient and the distinguished angle [Bleistein, 1987a,b]. Further, other geometrical attributes are determinable by processing additional slightly modified inversion formulas [Geoltrain and Chovet, 1991].

Thus, I opt for the principles underlying inversion—the wave equation and asymptotic theory—as opposed to the secondary principle underlying migration—that the ensemble satisfies the wave equation.

## REFERENCES.

- Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by the inversion of a causal Radon transform: *J. Math. Phys.*, **26**, 99-108.
- Bleistein, N., 1986, Two-and-one-half dimensional in-plane wave propagation: *Geophysical Prospecting*, **34**, 686-703.
- Bleistein, N., 1987a, On the imaging of reflectors in the earth: *Geophysics*, **52**, no. 7, 931-942.
- Bleistein, N., 1987b, Kirchhoff inversion for reflector imaging and sound speed and density variations, *in* Proceedings of EAEG/SEG Workshop on Deconvolution and Inversion, Rome, Italy.
- Bleistein, N., 1989, Large wave number aperture limited Fourier inversion and inverse scattering with appendices: *Wave Motion*, **11**, 113-136.
- Bleistein, N., Cohen, J. K., 1982, The Velocity Inversion Problem: Present Status, *New Directions: Geophysics* 47, 11, November, 1982.
- Bleistein, N., Cohen, J. K., and Hagin F., 1987, Two-and-one-half dimensional Born inversion with an arbitrary reference, *Geophysics* **52**, 26-36.
- Bleistein, N., Cohen, J. K., and Stockwell Jr., J., 1999, Mathematics of Multidimensional Seismic Inversion, Center for Wave Phenomena lecture notes, Colorado School of Mines, Golden, CO. Bleistein, N., and Gray, S.H., 1985, An extension of the Born inversion method to a depth dependent reference profile: *Geophysical Prospecting*, **33**, 999-1522.
- Clayton, R.W., and Stolt, R.H., 1981, A Born-WKBJ inversion method for acoustic reflection data: *Geophysics* 46, p. 1559-1567. (Reprinted in *Inversion of Geophysical Data: Geophysics Reprint Series No.9*, p. 52-60.)
- Cohen, J. K., and Bleistein, N., 1979, Velocity Inversion Procedure for Acoustic Waves: *Geophysics*, **4**, 6, 1077-1085.
- Cohen, J. K. and Hagin, F., 1985, Velocity inversion using a stratified reference, *Geophysics*, **50**, 1683-1700.

- Cohen, J. K., Hagin, F.G., and Bleistein, N., 1986, Three-dimensional Born inversion with an arbitrary reference: *Geophysics*, **51**, no. 8, 1552-1558.
- Docherty, P., 1987, Ray theoretical modeling, migration and inversion in two-and-one-half-dimensional layered acoustic media: Ph.D. thesis, Colorado School of Mines. Center for Wave Phenomena Research Report number CWP-051, Colorado School of Mines.
- Geoltrain, S., and Chovet, E., 1991, Automatic association of kinematic information to prestack images: Expanded Abstracts of the International Meeting and Exposition of the Society of Exploration Geophysicists, Tulsa, 890-892.
- Hagedoorn, J. G., 1954, A process of seismic reflection interpretation: *Geophysical Prospecting*, **2**, 85-127.
- Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions: *Geophysics*, **43**, 49-76.
- Stolt, R.H., and Weglein, A.B., 1985, Migration and inversion of seismic data: *Geophysics* 50, p. 2458-2472. (Reprinted in *Inversion of Geophysical Data: Geophysics Reprint Series No. 9*, p. 389-403.)
- Sullivan, M.F., and Cohen, J.K., 1987, Pre-stack Kirchhoff inversion of common offset data: *Geophysics*, **52**, no. 6, 745-754.

## Acknowledgement

I wish to express my gratitude to Sam Gray for making suggestions to improve the exposition and notation in this manuscript. Furthermore, the entire section entitled, “Returning to the string construction,” would not have been written were it not for his prodding. More generally, I wish to acknowledge Sam’s role over the years as *my own best critic*, always encouraging me to reach by exposing me to issues of practical implementation of my mathematical esoterica, when it would have been easier and safer for me not to do so. From my perspective, Sam has evolved from student to dear friend, respected colleague, peer—certainly—and sometimes master. In my role as a teacher, there can be no better reward than this.

## Appendix

Here is my own version of the magic of determining the envelope of a family of curves. First, we have to figure out how we are going to describe the family of curves. In the introductory discussion of the Hagedoorn construction, we used a parameter,  $\bar{x}$ , to characterize the center point of a family of circles and also to characterize the radii of the circles in equation (3). Earlier, in (1), we wrote down a more general form of that family of equations. Actually, even that equation has more detail than we need for this derivation, so let me simplify it, here. I suggest that a good starting point would be an equation of the form,

$$F(\mathbf{x}, \bar{x}) = 0, \tag{A-1}$$

with  $\mathbf{x} = (x_1, x_2)$  for this discussion. Comparing this equation with (1), one can see that the specific details of how the parameter  $\bar{x}$  describes the family of curves—the source position, the receiver position, the observed traveltime—have all been dispensed with in this simpler representation. For each  $\bar{x}$ , certain details of the function  $F$  change—we don't care how—to yield a different relationship between the two components of  $\mathbf{x}$ , hence a different curve.

Now, we are looking for another curve entirely, namely, the *envelope* of this family of curves. The first question to answer then is, “What is special about this curve that makes it an envelope of the family of curves in (A-1)?” Let us list the requirements:

1. Each point of this curve is on one of the curves of the family in (A-1).
2. At the contact of the envelope curve with an element of the family, the *tangent direction* to the envelope curve is the same as the tangent direction of the member of the family with which it is contact, up to a sign change.

The first requirement tells us that each point of the envelope is also a point on some curve of the family for some value of  $\bar{x}$ . That means that we can think of the envelope as a curve given parametrically, with  $\bar{x}$  as the parameter. That is,

$$\mathbf{x} = \mathbf{x}(\bar{x}) \tag{A-2}$$

and

$$F(\mathbf{x}(\bar{x}), \bar{x}) \equiv 0. \tag{A-3}$$

The emphatic equal sign here means that, if we actually had the explicit expressions for  $\mathbf{x}(\bar{x})$  in hand and substituted them into  $F$ , the result on the left would visibly be zero; this equality does not hold only for some choice of  $\bar{x}$ , but for *every* choice of  $\bar{x}$ .

Identities such as this one are really nice mathematical tools, because no matter how many times we differentiate them, we always get zero. On the other hand, because the differentiation on the left side involves using the chain rule, we learn a good deal about the interplay of the derivatives of  $F$  with the derivatives of the parametric functions that describe the envelope in terms of  $\bar{x}$ . So, let us now differentiate this identity in  $\bar{x}$  with the help of the chain rule:

$$\left[ \frac{\partial F}{\partial x_1} \frac{dx_1}{d\bar{x}} + \frac{\partial F}{\partial x_2} \frac{dx_2}{d\bar{x}} \right] + \frac{\partial F}{\partial \bar{x}} = 0.$$

The expression in square brackets is a dot product of the gradient of the function,  $F$ , with the tangent to the envelope curve,  $\mathbf{x} = \mathbf{x}(\bar{x})$ . For each curve in the original family—that is, for each  $\bar{x}$ —the gradient of  $F$  is orthogonal to the curve,  $F = 0$ , for every choice of  $\mathbf{x}$  on the curve. More precisely, that means that, for each curve of the family (each fixed  $\bar{x}$ ), this gradient is perpendicular to the tangent to the curve  $F(\mathbf{x}, \bar{x}) = 0$ .

Now, we are ready to invoke condition 2 above. Right where the curve,  $F = 0$  is in contact with the envelope  $\mathbf{x} = \mathbf{x}(\bar{x})$ , *the tangents of the two curves must be colinear or anti-colinear!* Thus, this dot product is zero, right where the envelope curve touches a member

of the original family of curves and this last equation reduces to

$$\frac{\partial F(\mathbf{x}, \bar{x})}{\partial \bar{x}} = 0. \tag{A-4}$$

Now look at this last equation and (A-1) together. This is a pair of equations in  $\mathbf{x}$  and  $\bar{x}$  that must be satisfied for points on the envelope of the family of curves. These are two equations in three unknowns that characterize the envelope curve. These are the standard equations derived in calculus texts for the determination of the envelope of a family of curves. Unfortunately, the derivation usually precedes a discussion of parametric equations and gradients. Thus, the explanation given here in terms of orthogonality of a gradient and a tangent vector is usually done with a bit of handwaving, which, I believe, contributes to the confusion about this whole subject.

We now have two ways to proceed. One way is to solve for two of the unknowns, say,  $\mathbf{x}$ , in terms of the third unknown,  $\bar{x}$ . Doing so, would give us the parametric representation of the envelope that we have denoted by the function,  $\mathbf{x}(\bar{x})$ . The other way of solving the problem is to solve one of the equations for  $\bar{x}$ , usually the second, as a function of  $\mathbf{x}$ , and then substituting the result back into the other equation, say, (A-1). Then, we obtain an equation totally in  $\mathbf{x}$ ,

$$F(\mathbf{x}, \bar{x}(\mathbf{x})) = 0. \tag{A-5}$$

This is an equation of a single curve having the properties listed above that were necessary for an envelope of a family of curves. Any chicanery to solve for  $\bar{x}$  between *both* equations, is perfectly acceptable and substitution into *either* equation will produce a valid description of the envelope.

To review: the envelope of a family of curves  $F(\mathbf{x}, \bar{x}) = 0$  is determined by setting the first derivative with respect to  $\bar{x}$  equal to zero, and then deriving either a parametric representation of the envelope from this pair of equations— $F = 0$  and  $\partial F/\partial \bar{x} = 0$ —or solving for  $\bar{x}$  and substituting into one of the equations.