Let us consider the forward modeling choice for λ. At 30 Hz, 500 m and a propagation speed of 2000 m/sec, $\lambda = 47$! At the same 30 Hz, 5000 m and a propagation speed of 6000 m/sec, λ is about three times larger. So, if we use the range to a reflector as the length scale, λ is well into the range in which 1% accuracy is achieved.

On the other hand, it is also necessary that individual reflectors be "well separated" by this scaling criterion. So, let us consider two reflectors, 100 m apart, at the same 30 Hz, with a propagation speed of 6000 m/sec. Then, $\lambda = \pi$. This last example is more realistic. It reflects the limitation of interpretations based on the method of stationary phase with isolated stationary points. Note that the separate parameter values here are about at the limits of what one would expect from experience. Our last plot suggests that—for the Hankel function, at least—the single term asymptotic expansion provides an estimate of the exact solution with an error of about 4% for this last set of parameters.

Of course, near pinch-outs where the length scale (and, hence, λ) approach zero, the simple method of stationary phase will fail. Indeed, there are real problems with amplitude accuracy right at the edges of reflectors, such as might occur along a fault. Further, near buried foci, the method also fails for more subtle reasons. So, it becomes more a matter of learning when the method is valid and, even more importantly, when it is not. For a large class of seismic implementations, the method is virtually exact. In fact, most interpretations of wave propagation and imaging are based on concepts derived from plane wave propagation in the hope that things don't change too much when the wave fronts and/or the reflectors are curved (a "little" bit) and the medium parameters do not vary "too much" over a few wavelengths. With some quantification, those qualitative ideas provide a basis for parameter bounds that make the method of stationary phase the method of choice for analyzing and interpreting integral processing schemes for seismic data processing.

Transforming Hagedoorn's construction into a migration program.

The Hagedoorn process provides guidance for the development of a computer program that would produce an image of the subsurface. On each trace, a particular data point—a particular traveltime—was used to draw a circular isochron. In drawing that isochron, we are simply distributing the data from a given time on a given trace over the equi-traveltime curve representing the possible locations of reflection points. Constructive interference then produced the image as the envelope of isochrons.

Now, let us consider doing the same thing with a computer code. The problem is, we would not want to search each trace for the "significant" data points that should be spread over isochrons. Instead, let us take *all* points on the trace, take the *given data value* at each trace point and distribute the data over all bins representing points on the appropriate isochrons in the image domain. We would then be relying on constructive and destructive interference of values in any particular bin to give us a "strong" response in some bins and a weak response in others. Also, note that in our shoestring construction, we *chose* traveltimes at which to draw isochrons. Now, *all* traveltimes are candidates and the data values at those times act as a measure of how important that contribution might be to the overall picture.

Indeed, this process will work, as long as we take care of some minor scaling and filtering issues. For example, it is reasonable to expect that signals at later times would be weaker than signals at earlier times. Hence, in order that the strength of images at different depths have comparable intensity, we might think to scale up the data by some multiple and some power of t before distributing it over the appropriate bins. Somewhat more subtle might be the need to do some frequency domain filtering on the data so that the images appear as bandlimited delta functions rather than doublets of one sort or another. (The frequency domain source signatures, $i\omega$, $\text{sgn}(\omega)$, $\sqrt{|\omega|}\exp\{i\pi\,\text{sgn}(\omega)/4\}$, all produce doublets in the time domain and all arise in the analysis of Kirchhoff migration and inversion algorithms in various dimensions.) In any case, these are issues that I will not deal with, at the moment.

Now, let us think about turning the processing around in the following sense. Instead of distributing the data from each point on a trace—say, each fixed $(\bar{x}, t)$—over an array of output points, $\boldsymbol{x} = (x_1, x_3)$, let us think about fixing $\boldsymbol{x}$ and asking the following question, "On each trace (each $\bar{x}$), at what time (which t) would data arrive at $\bar{x}$ after having traveled from there to $\boldsymbol{x}$ and back again?" When we know that, we would merely add data to the bin at $\boldsymbol{x}$ from each such trace at each such time. Thus, in the former method, the trace point, $(\bar{x}, t)$, is fixed and the data are distributed along a curve in the image domain, that is, among $\boldsymbol{x}$-values. In the latter method, the point, $\boldsymbol{x}$, is fixed and data are *collected* from an ensemble of points in the $(\bar{x}, t)$-domain.

Of course, in analogy with the circles of Figure 2, this family of trace/traveltimes also fits on a set of nice tidy curve, but now in the $(\bar{x}, t)$ domain. For our simple zero offset constant wavespeed experiment, it is just equation (3), however, recast in a form that reveals the nature of the curve in the $(\bar{x}, t)$ domain, namely,

$$\frac{t^2}{(2/c)^2} - (\bar{x} - x_1)^2 = x_3^2. \tag{9}$$

Of course, this is just the familiar diffraction hyperbola of migration.

The generalization of this simple case suggested by equation (1) tells us how to find the traveltime on each trace in the most general of problems. In that equation, the source and receiver positions are both described as functions of $\bar{x}$. For example, for a common offset experiment on a flat datum surface,

$$\boldsymbol{x}_s(\bar{x}) = (\bar{x} - h, 0), \quad \boldsymbol{x}_r(\bar{x}) = (\bar{x} + h, 0),$$

with zero offset just being the special case of this equation for which $h = 0$. As a second example, for a common source experiment on a flat datum surface,

$$\boldsymbol{x}_s(\bar{x}) = (\text{constant}, 0), \quad \boldsymbol{x}_r(\bar{x}) = (\bar{x}, 0).$$

Then, finding $t(\bar{x})$ amounts to carrying out the necessary ray tracing to determine the sum of traveltimes from source and receiver to output point. Once that is done, an integration (summation) over the data values at those $(\bar{x}, t(\bar{x}))$ coordinates can be carried out for each point $\boldsymbol{x}$ in the image domain to create an output at that point. As with the shoestring construction, we would rely on constructive and destructive interference to produce an image out of all these summations.

Again, we must invoke our caveat that some weighting and frequency domain filtering might be necessary. Modulo knowing those weights, though, I propose that we have a basis for writing down an *integration formula* for imaging data. That formula, familiar to Kirchhoff migrators and inverters, is

$$\beta(\boldsymbol{x}) = \int d\bar{x}\ B(\boldsymbol{x}, \bar{x}) D(\bar{x}, \tau(\boldsymbol{x}, \bar{x})). \tag{10}$$

In this equation, $D(\bar{x}, \tau(\boldsymbol{x}, \bar{x}))$ is the Fourier filtered data on the trace labeled by $\bar{x}$, its first argument, and evaluated at the traveltime from source to output point to receiver, that we have denoted by $\tau(\boldsymbol{x}, \bar{x})$ as a simplification of the notation for traveltime, $\tau(\boldsymbol{x}, \boldsymbol{x}_r(\bar{x}), \boldsymbol{x}_s(\bar{x}))$. There will be more discussion of the Fourier filtering, below.

The effect of the integration over $\bar{x}$ is to carry out the summation over all of the traces, with the filtered data on each trace picked out at the appropriate time on each trace. On each trace, the time $t(\bar{x})$ is chosen to be the traveltime, $\tau(\boldsymbol{x}, \boldsymbol{x}_r(\bar{x}), \boldsymbol{x}_s(\bar{x}))$. Of course, we have allowed for a spatial weighting factor, $B(\boldsymbol{x}, \bar{x})$, that could depend both on the trace location and on the output location. Also, in a subtle way, we have allowed for an acquisition *curve*, rather than the acquisition line of our examples. All we need to do is express that curve parametrically through the functions, $\boldsymbol{x}_r(\bar{x})$ and $\boldsymbol{x}_s(\bar{x})$.

This is the fundamental form of Kirchhoff migration/inversion. We have seen, here, how the formula follows from a thoughtful examination of Hagedoorn's construction. Of course, for this generalization, migration on a shoestring is no longer possible and there is a whole art and science built around the issues of carrying out the spatial integration at minimal cpu. However, the fundamental structure of this formula is inescapable.

This form also leads the way to a 3D migration/inversion formula. The only difference is that, now, the sources and receivers will be distributed over a suface, rather than over a line or curve. For this extension, then, we need to describe the acquisition geometry by a function of two variables, $\bar{\boldsymbol{x}} = (\bar{x}_1, \bar{x}_2)$. Then,

$$\boldsymbol{x}_s = \boldsymbol{x}_s(\bar{\boldsymbol{x}}) = (x_{s1}(\bar{\boldsymbol{x}}), x_{s2}(\bar{\boldsymbol{x}}), x_{s3}(\bar{\boldsymbol{x}})), \quad \boldsymbol{x}_r = \boldsymbol{x}_r(\bar{\boldsymbol{x}}) = (x_{r1}(\bar{\boldsymbol{x}}), x_{r2}(\bar{\boldsymbol{x}}), x_{r3}(\bar{\boldsymbol{x}})),$$

and (10) is replaced by

$$\beta(\boldsymbol{x}) = \int d\bar{x}_1 d\bar{x}_2\ B(\boldsymbol{x}, \bar{\boldsymbol{x}}) D(\bar{\boldsymbol{x}}, \tau(\boldsymbol{x}, \bar{\boldsymbol{x}})). \tag{11}$$

Now, the output point is $\boldsymbol{x} = (x_1, x_2, x_3)$.

Any reader who has done their own wordprocessing can recognize that this new result is merely a copy and past of the 2D result, (10), with some minor editing. Of course, the devil is in the details of the functions, B and D in these two formulas, but the stucture remains unchanged.

Returning to the string construction

Hagedoorn's string construction is not lost in the results (10) and (11); it is just hiding. It is necessary to carry out these computations for a whole domain of output points $\boldsymbol{x}$. Let us consider doing all of those at once. To do so, proceed as follows.

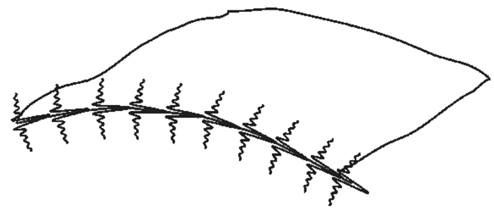

FIG. 5. The (bandlimited) singular function of a surface. The singular function is a delta function of a single argument that is zero on the surface.

1. Choose a particular time on a particular data trace; that is, a particular choice of $\bar{x}$ in (10) or $\bar{\boldsymbol{x}}$ in (11), and a corresponding $t(\bar{x})$ or $t(\bar{\boldsymbol{x}})$, respectively.

2. Determine the $\boldsymbol{x}$-values of the isochron defined by (1), or its natural extension to 3D.

3. For each such $\boldsymbol{x}$, compute the weighted contribution of that data point to the integration at each point $\boldsymbol{x}$ on the isochron in the image domain.

4. Repeat for all data points on all traces.

In carrying out this process, we are distributing the data at each data point on a trace over its isochron back in the physical space. That constitutes a return to Hagedoorn's method, but in the context of the complete generalization to variable background wavespeed and source/receiver configuration proposed here. Furthermore, where in our shoestring construction, we could choose the few points on the trace that we wanted to distribute over the image domain, here, we treat every point on every trace as a potential contributor and allow the amplitude of the data at each such point reflect the significance of that data point to the overall image.

Turning Kirchhoff migration into Kirchhoff inversion.

The question that now arises is how to choose the frequency domain filter and spatial weight to turn (10) or (11) into an inversion formula. In order to decide that, we must first think about what we might have a right to expect from such a formula. Note, first, that it is based on a string construction applied to reflection data from a single reflector. The very term, "reflection data" implies high frequency asymptotics: for Laplace's equation—$\omega = 0$—there are no reflections and, thus, for small values of ω, the solution is just a power series in ω with the solution of Laplace's equation as the first term; solutions of Laplace's equation do not exhibit reflection. Thus, I would propose that we might have a right to expect this formula to produce an image for a single reflector with an *amplitude* that was meaningful with respect to the reflection strength at the image points on the reflector only as a high frequency asymptotic approximation. Anything more than that would be icing on the cake.

In fact, before the approach that I have proposed here was developed, the asymptotic inversion of constant background zero offset data had already been achieved [Cohen and Bleistein, 1979]. We were able to analyze the application of that method to reflection data, asymptotically. The output was proportional to a bandlimited delta function that peaked