Ground Roll and Radial Traces

Jon F. Claerbout

The land profile in figure 1 contains ground roll that is spatially aliased. Evidence is the appearance of reverse moveout in the slow waves. This data is worth studying both for the ground roll and for other reasons. In geophysics we generally lack adequate spatial coverage in one way or another, yet most of our analytic techniques demand that we have nonaliased data. The dataset in figure 1 seems to be a clean superposition of regular reflections and regular ground roll. Although the data is clearly aliased, the simplicity of the underlying model encourages belief that the aliasing should not frustrate analysis. A useful goal is to decompose this data into its reflection component and its ground-roll component. The immediate value would be an improved stack. Longer range value could arise from applying techniques learned here to our many other problems with aliasing.

Besides the many well known advantages of the radial trace coordinate system, it also offers some special properties in the analysis of ground roll. Since ideal ground roll has a travel time simply proportional to distance, the radial $r = x/t$ axis nicely organizes the ground roll by velocity. Something useful happens on the time axis too. *Idealized ground roll is transformed to zero temporal frequency by the radial trace transformation.* Analogously, moveout correction transforms reflections to the zero spatial frequency.

Besides the utility of the radial trace coordinate system in the analysis of ground roll, it is a natural coordinate system to use to study angle-dependent phenomena. Those of you who keep your eye near the the oil patch will recall that the distinction between productive bright spots and nonproductive ones can sometimes be made on the basis of angle dependence of the reflectivity. It seems (ref - Ostrander's 1982 SEG talk.) that the heart of the problem is being able to calibrate the background signals, and for this radial coordinates seem ideal. I would like to see some fundamental studies on the radial dependence of seismic spectra as well as radial dependence of the spherical divergence correction.

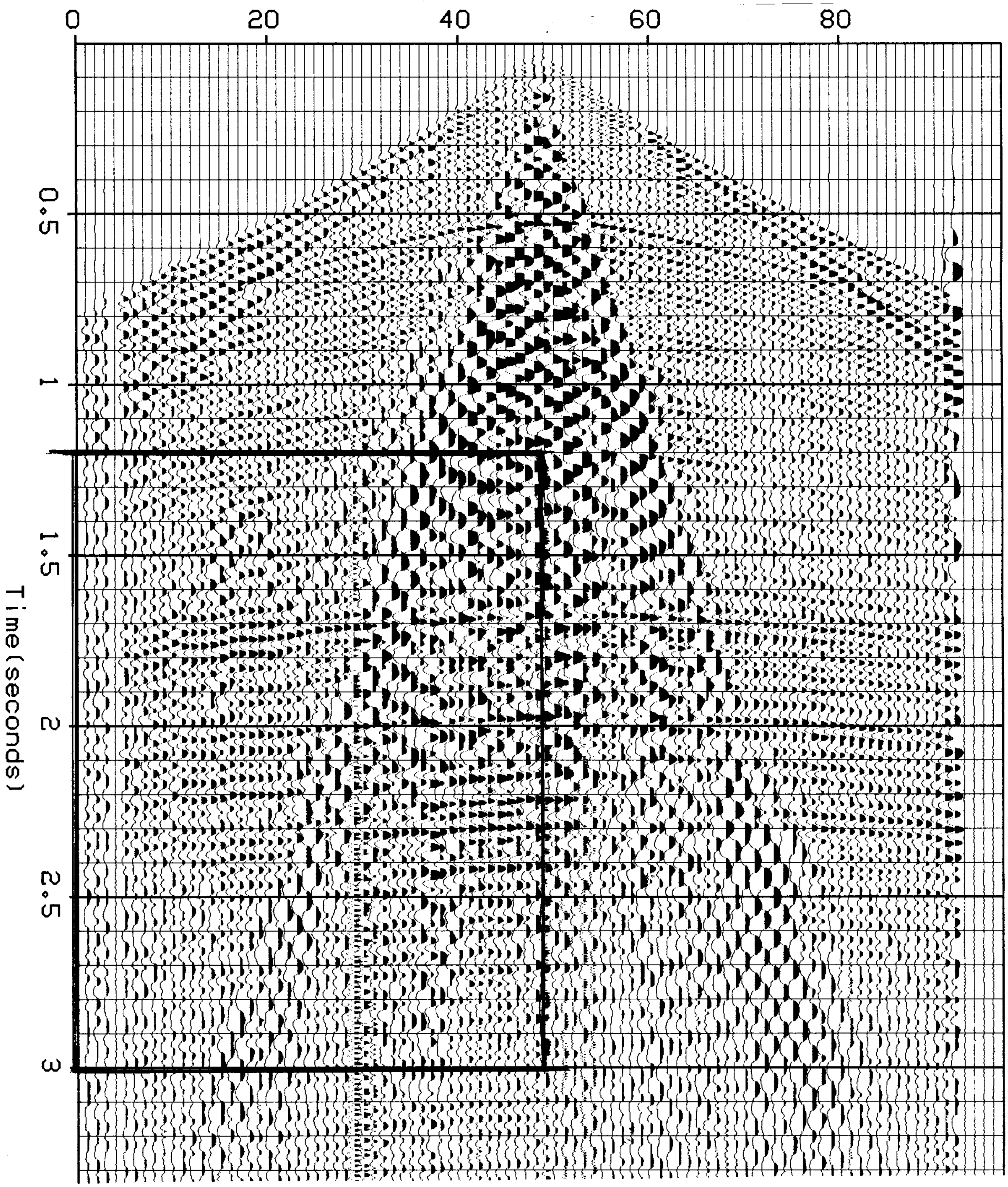

FIG. 1. Land profile from North America with notable ground roll. (Western Geophysical)

Interpolation of the Data

I began my efforts with the interpolation of the data, that is, with estimating seismograms inbetween the observed ones. This seems to be the part of the problem that is intractible by the classic definition of aliasing. Figure 2 shows a magnified portion of figure 1 with several interlaced zero traces. To ensure visibility of the data in all locations, the dynamic range was compressed by two means: All traces were divergence corrected by multiplying by a power of time (A typical plot was scaled by $t^{2.5}$). Second, each point was replaced by its signed square root.

I tried many interpolation methods. A typical result is shown in figure 3. As the deadline of this progress report arrived, I found myself with many workable methods but none that seemed ideal. Perhaps the reason that so many methods produced generally pleasing results was that radial trace organization was used for all methods. On the other hand, none of the methods seemed to be ideal. When studying the interpolations, I used Rick Ottolini's interactive movie program which allowed me to flicker back and forth between various frames and various gain schemes. The interpolations were never so good that I had trouble distinguishing the real from the interpolated data. However, I often found particular display parameters on which it was difficult to distinguish real from interpolated data. For example, taking the square root to compress the dynamic range also happens to mask the fact that interpolated traces tend to be of lower amplitude. In trying to design a method which measures statistical characteristics of the data, it is a problem that whatever statistical measures are taken, they change with location on the data plane.

Before I launch into a long-winded theoretical analysis, I should report that visually pleasing interpolations arose from even such simple ideas as adding the linear interpolation in the time direction to the linear interpolation along the space direction. But theoretically, even this is not straightforward. The first question is how the two interpolations should be averaged. Where the ground roll dominates it would seem that the time interpolation should be more strongly weighted. Where the reflections dominate it would seem that the space interpolation should be more strongly weighted. We are quickly led into statistical models for the data itself.

Removing the Missing Data That Fits No Propagation Mode

First a general multidimensional missing data algorithm will be developed, then it will be specialized to the ground roll problem. Let $\mathbf{F}_j$ transform a data set so that some mode of propagation becomes a low frequency in some direction. (It could be a VSP with $\mathbf{F}$ being the shift operator.) Let $\mathbf{L}$ denote a weight which selects a narrow band of low frequencies (in

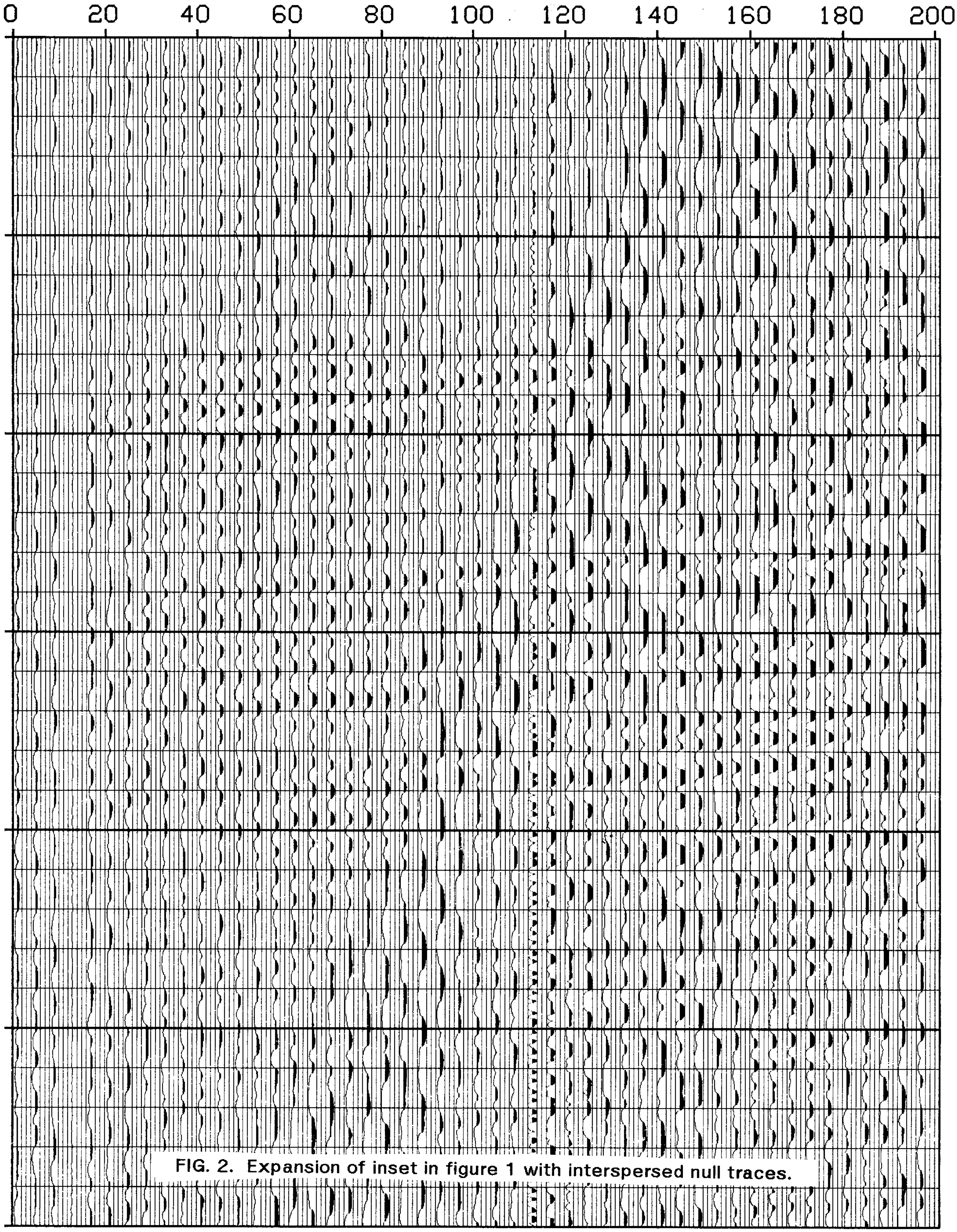

FIG. 2. Expansion of inset in figure 1 with interspersed null traces.

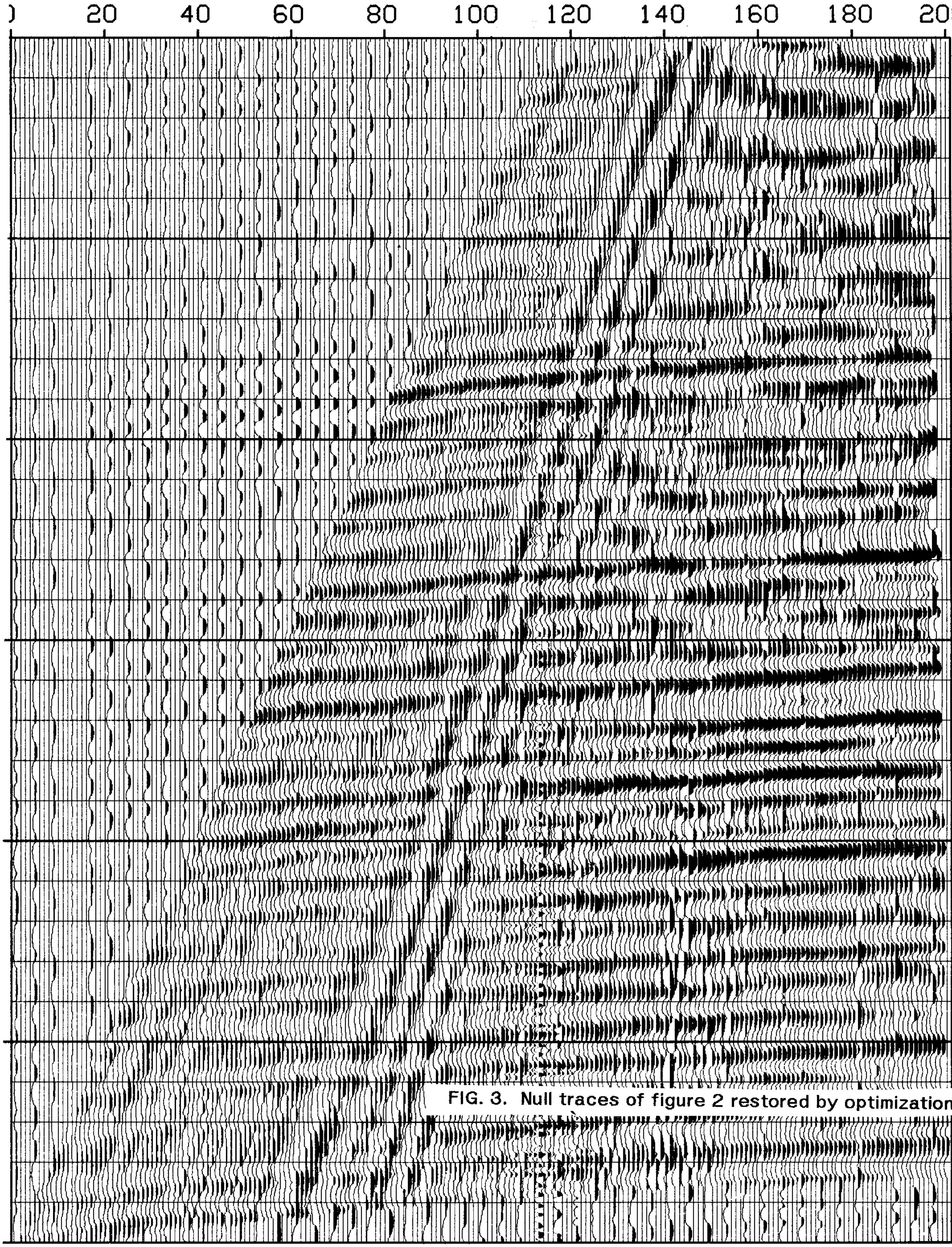

FIG. 3. Null traces of figure 2 restored by optimization.

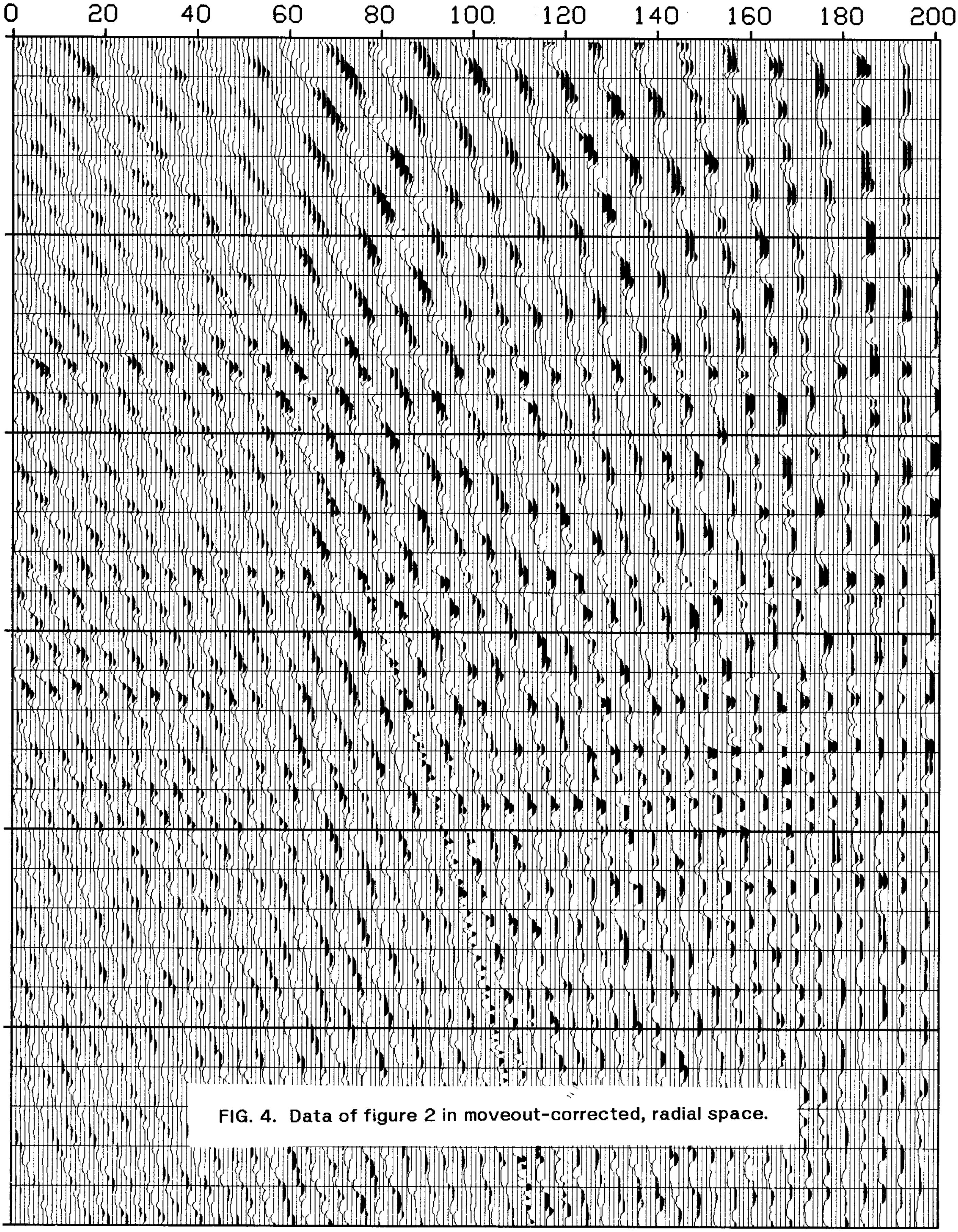

FIG. 4. Data of figure 2 in moveout-corrected, radial space.

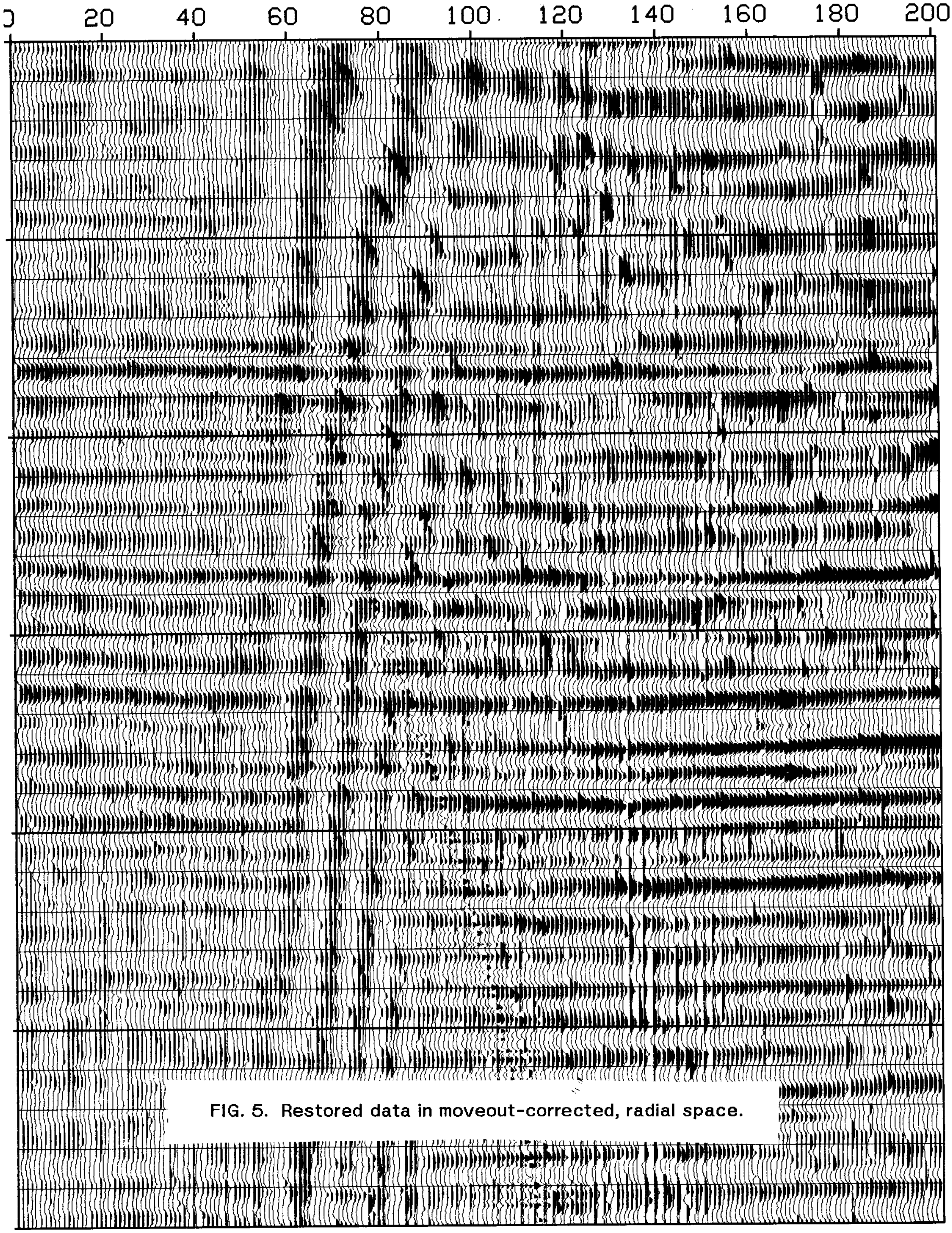

FIG. 5. Restored data in moveout-corrected, radial space.

that direction) and $\mathbf{H} = 1 - L$ a weight that selects the wide band of remaining frequencies. Data does not fit the propagation mode if the operator and data are such that $|F_j^{-1}\mathbf{H}\mathbf{F}_j\mathbf{d}| \approx |d|$. Data fits the mode well if $|F_j^{-1}\mathbf{H}\mathbf{F}_j\mathbf{d}| \ll |d|$. Think of

$$\mathbf{H}_j \; = \; \mathbf{F}_j^{-1}\mathbf{H}\mathbf{F}_j \; = \; No_j \; = \; 1 \; - \; Yes_j \tag{1}$$

Missing data should be subtracted out from itself if it fits no mode of propagation. This compounding of negatives can be thought of as

$$\mathbf{d} \; = \; \mathbf{d} - \mathbf{M}(1 - Yes_1)(1 - Yes_2) \cdots (1 - Yes_n)\mathbf{d} \tag{2}$$

More specifically, the missing data iteration is

$$\mathbf{d} \; = \; \mathbf{d} - \mathbf{M}(\mathbf{F}_1^{-1}\mathbf{H}\mathbf{F}_1)(\mathbf{F}_2^{-1}\mathbf{H}\mathbf{F}_2) \cdots (\mathbf{F}_n^{-1}\mathbf{H}\mathbf{F}_n)\mathbf{d} \tag{3}$$

Now let us specialize to the case of ground roll in moveout-corrected, radial-trace coordinates. The transformations $\mathbf{F}_j$ are unnecessary, we can just apply one highpass filter along the r-axis and the other along the t-axis, say

$$\mathbf{d} \; = \; \mathbf{d} - \mathbf{M}\,\mathbf{H}_r\,\mathbf{H}_t\,\mathbf{d} \tag{4}$$

In computing the figures, I ran four iterations. This took a few minutes on the VAX with nothing optimized. Actually I solved for a scaling factor that scaled **M**, but I found it always turned out very close to the stability limit 2.0. I will report further on the solution method after I have greater confidence that I know what to minimize. Further iterations did change the solution, but not substantially.

The Implied Two-Dimensional Spectrum

The iteration (4) is known to solve a least squares problem. In fact $\mathbf{H}_x\,\mathbf{H}_t$ can be interpreted as the inverse of the covariance matrix of the data **d**. This means that the method presumes a data spectrum

$$E[\mathbf{d}\,\mathbf{d}^T] \; = \; \frac{k_0^2 + k^2}{k^2}\,\frac{\omega_0^2 + \omega^2}{\omega^2} \tag{5}$$

There are two adjustable bandwidth parameters in (5). I experimented with them. Subjectively, the results did not seem to be very sensitive to the cutoffs. The spectrum (5) may not be exactly your idea of the spectrum that the data should have. But at least it is a separable function, and one could hope that the detailed functional form is not essential to making an improvement over conventional approaches.

Separability Difficulty in the Statistical Formulation

Being familiar with wave equation migration in laterally variable media, we are all skilled in manipulating tridiagonal matrices. In statistical analysis, tridiagonal matrices can be useful for representing covariance. There are two main degrees of freedom in a tridiagonal matrix, first a spectral parameter, the spatial correlation length and second, a time domain parameter, the envelope. Let e_g denote the envelope function of the *g*round roll. Let e_r denote the envelope function of the *r*eflectors. These envelopes vary with time and space. Let $\mathbf{H}_g$ denote the temporal high-pass filter that tends to whiten the temporal spectrum of *g*round roll. Let $\mathbf{H}_r$ denote the spatial high-pass filter that tends to whiten the spatial spectrum of moved out *r*eflectors. These filters are made up in the usual way with tridiagonal matrix representations of the second difference operator. The filters allow us to estimate the envelopes from the data. I propose to represent the spectrum of the data as the spectrum of the ground roll plus the spectrum of the reflectors:

$$S \quad = \quad \frac{e_g}{\mathbf{H}_g} + \frac{e_r}{\mathbf{H}_r} \tag{6}$$

The inverse of the spectrum is what we need in missing data iteration.

$$\frac{1}{S} \quad = \quad \frac{1}{\dfrac{e_g}{\mathbf{H}_g} + \dfrac{e_r}{\mathbf{H}_r}} \tag{7}$$

The implied missing data iteration is:

$$\mathbf{d} \quad = \quad \mathbf{d} \; - \; \mathbf{M} \left[\frac{1}{\dfrac{e_g}{\mathbf{H}_g} + \dfrac{e_r}{\mathbf{H}_r}} \mathbf{d} \right] \tag{8}$$

Unfortunately the quantity in the square brackets is not computationally practical because the operators $\mathbf{H}_r$ and $\mathbf{H}_g$ operate in different directions. So I considered approximations. Equation (7) is exactly equal to

$$\frac{1}{S} \quad = \quad \frac{\mathbf{H}_r \, \mathbf{H}_g}{e_g \, \mathbf{H}_r \; + \; e_r \, \mathbf{H}_g} \tag{9}$$

Whenever a denominator gives problems the safe course seems to be to over-estimate it, so we can try a denominator approximation, say replace the denominator in (9) by $e_{\mathbf{d}} = e_r + e_g$.

$$\frac{1}{S} \quad \approx \quad \frac{\mathbf{H}_r \, \mathbf{H}_g}{e_{\mathbf{d}}} \tag{10}$$

This you will recognize is close to the the spectrum implied by iteration equation (4) which was used to generate the figures.

Of course $e_{\mathbf{d}}$ is just the data envelope, so we have lost whatever advantages are associated with distinguishing between the envelope of ground roll and that of the reflections. I am afraid that this is a serious loss, and it may account for a problem in the results. Notice in figure 3 that there seems to be too much ground roll at wide angles. If the estimation procedure were making use of e_r / e_g we might not have that problem.

The Algorithm

When I developed this algorithm I thought it might be vintage steepest descent or conjugate direction, but when I looked in Luenberger's book on optimization, it seems that he does not have anything like the constraints that we require to match the field data exactly. Let the data at the nth iteration be a sum of past directions g_i times past distances α_i.

$$\mathbf{d}_n \quad = \quad \mathbf{d}_0 + \sum_{i=1}^{n} \alpha_i \, \mathbf{g}_i \tag{11}$$

Let the directions be chosen from

$$\mathbf{g}_{n+1} \quad = \quad -\,\mathbf{M}\,\mathbf{H}_r\,\mathbf{H}_g\,\mathbf{d}_n \tag{12}$$

$$\mathbf{g}_{n+1} \quad = \quad -\,\mathbf{M}\,\mathbf{H}_r\,\mathbf{H}_g \left(\mathbf{d}_0 + \sum_{i=1}^{n} \alpha_i \, \mathbf{g}_i \right)$$

$$\mathbf{g}_{n+1} - \mathbf{g}_n \quad = \quad -\,\mathbf{M}\,\mathbf{H}_r\,\mathbf{H}_g\,\alpha_n\,\mathbf{g}_n \tag{13}$$

Define a new direction vector $\mathbf{v}_n$

$$\mathbf{v}_n \quad = \quad \frac{\mathbf{g}_{n+1} - \mathbf{g}_n}{\alpha_n} \quad = \quad -\,\mathbf{M}\,\mathbf{H}_r\,\mathbf{H}_g\,\mathbf{g}_n \tag{14}$$

Now define a quadratic form to be minimized.

$$Q(\alpha_n) \quad = \quad (\mathbf{d}_{n-1}^T + \mathbf{g}_n^T \alpha_n)\, \frac{\mathbf{H}_r\,\mathbf{H}_g}{e} (\mathbf{d}_{n-1} + \alpha\,\mathbf{g}_n) \tag{15}$$

In this expression the envelope of the data appears as a divisor. Technically, it should be represented as a diagonal matrix. Notice that the direction chosen by equation (12) doesn't have the envelope division and it is not the same as the gradient of the quadratic form (15). This is because age and craftiness has taught us not to descend with the gradient, but rather in a direction scaled to the data like a highpass filter. Then the change in the data is proportional to the data. Setting to zero the derivative of Q with respect to α_n gives

$$\alpha_n \quad = \quad - \frac{g_n^T \, \frac{\mathbf{H}_r \, \mathbf{H}_g}{e} \mathbf{d}_{n-1}}{g_n^T \, \frac{\mathbf{H}_r \, \mathbf{H}_g}{e} g_n} \quad = \quad - \frac{-g_n^T \, \frac{1}{e} g_n}{-g_n^T \, \frac{1}{e} \mathbf{v}_n} \tag{16}$$

I implemented this in a program like:

$\mathbf{d}_0 \quad = \quad \mathbf{K}\mathbf{d}$

$\mathbf{g}_1 \quad = \quad -\mathbf{M}\,\mathbf{H}_r\,\mathbf{H}_g\,\mathbf{d}_0$

for $i = 1, n$ {

$\quad \mathbf{v}_i \quad = \quad -\mathbf{M}\,\mathbf{H}_r\,\mathbf{H}_g\,\mathbf{g}_i$

$$\alpha \quad = \quad - \frac{\mathbf{g}_i^T \, \frac{1}{e} \mathbf{g}_i}{\mathbf{g}_i^T \, \frac{1}{e} \mathbf{v}_i}$$

$\quad \mathbf{d}_i \quad = \quad \mathbf{d}_{i-1} + \alpha\,\mathbf{g}_i$

$\quad \mathbf{g}_{i+1} \quad = \quad \mathbf{g}_i + \alpha\,\mathbf{v}_i$

}

It turned out that the step size, α was always about equal 2.0 which is a like a stability limit. In fact, it was so close to 2.0, always in a range say 1.8-2.2, that I can imagine not bothering to compute it.

Geographical distribution: The Class of 1985, Harvard and Stanford