Interpolating Aliased Seismic Sections

Bert Jacobs

Abstract

A straight line model for the power spectrum of a seismic section leads to an iterative algorithm for simultaneously unscrambling the aliasing and interpolating. Sections were used because a seismic section with some high dips on it resembles a 3-D data set in the cross-line direction. Interpolation of such data sets is an important unsolved problem in geophysical data processing.

Introduction

Deciding which of several aliases present in a section is the correct one will require a model for seismic sections. One useful model is that a section is made up of a collection of linear events. Thus, one "popular" method of interpolating at a point in a section is to scan in dip space and to interpolate linearly along the dip of maximum coherence. This leaves begging the question of what to do when the dip spectrum has several peaks.

Geophysicists often approach problems which are difficult in the space-time domain in one of several frequency domains. Since the power spectrum of a line in x,t-space is an amplituded modulated line in k,ω-space, it seems appropriate to examine the aliasing/interpolation problem in the FK domain.

An estimate of the de-aliased power spectrum can be obtained by smoothing in radial directions in the FK domain. Given such an estimate the various aliases present can be untangled. We will restrict attention here to data which is two-fold aliased in the k-direction only.

Removing Aliasing and Interpolating in One Dimension

Aliasing is an artifact of sampling. If high frequency data is sampled at a low rate and Fourier transformed, the transform will have energy at unexpected frequencies. Consider a one-dimensional, causal, real, and continuous time function f(t) which we sample at rate $1/\Delta t$. The discrete Fourier transform (DFT) of this time series is given by

$$F(\omega) = \sum_{m=0}^{N-1} f(m\Delta t) e^{-2m\pi\omega\Delta t}$$

The continous Fourier transform (FT) of the series is more fundamental than the DFT

$$\hat{F}(\omega) = \int_0^\infty dt\ f(t)\ e^{-2\pi\omega t}$$

Both F and $\hat{F}$ are Hermitian functions of ω. If $\hat{F}(\omega) = 0$ for $|\omega| > 1/2\Delta t$ then

$$F(\omega) = \hat{F}(\omega) \qquad (0 < \omega < 1/2\Delta t)$$

$$F(\omega) = \hat{F}(\omega - 1/\Delta t) \qquad (1/2\Delta t < \omega < 1/\Delta t)$$

When $\hat{F}(\omega) \neq 0$ for $1/2\Delta t < |\omega| < 1/\Delta t$ and $\hat{F}(\omega) = 0$ for $|\omega| > 1/\Delta t$ this simple situation changes to

$$F(\omega) = \hat{F}(\omega) + \hat{F}(\omega - 2/\Delta t) \qquad (0 < \omega < 1/2\Delta t) \tag{1}$$

$$F(\omega) = \hat{F}(\omega - 1/\Delta t) + \hat{F}(\omega) \qquad (1/2\Delta t < \omega < 1/\Delta t) \tag{2}$$

The effect of two-fold aliasing is to additively mix two frequency components in each DFT component. This mixing must be untangled if a proper job of interpolation is to be done.

"De-aliasing" and interpolation can be done simultaneously if an estimate of the output power spectrum is at hand. Suppose we have the real fast Fourier transform (RFFT) of a real time series, N points long. Their will then be $N/2$ positive frequency components available. The interpolation problem in the time domain will correspond to an extrapolation problem in the frequency domain. Negative frequencies (the last $N/2$ points) need not be considered because of symmetry.

$$F_k \rightarrow \textit{input transform} \qquad k = 0,1,2,\cdots,N/2$$

$$G_k \rightarrow \textit{output transform} \qquad k = 0,1,2,\cdots,N \tag{3}$$

$$|H_k| \rightarrow \textit{desired spectrum} \qquad k = 0,1,2,\cdots,N$$

If we assume that their is only two-fold aliasing then

$$G_k + G_{k+N/2} = F_k \qquad k = 0,1,\cdots,N/2 \tag{4}$$

$$|G_k| = |H_k| \qquad k = 0,1,...,N \tag{5}$$

Substituting the equation (4) into equation (5), and defining φ_k to be the phase of G_k, yields N equations in N real unknowns

$$|H_k|\cos\varphi_k + |H_{k+N/2}|\cos\varphi_{k+N/2} = \text{Re } F_k \qquad k = 0,1,\cdots,N/2 \tag{6}$$

$$|H_k|\sin\varphi_k + |H_{k+N/2}|\sin\varphi_{k+N/2} = \text{Im } F_k \qquad k = 0,1,\cdots,N/2 \tag{7}$$

which may or may not have a reasonable solution.

The lesson here is that a good model of the power spectrum (equation (5)) goes a long way towards solving the interpolation problem. The rest of the way is traversed with equation (4), which is constraint that the input data be honored in choosing the phase spectrum. These two principals will govern our choice of an algorithm for interpolating and "de-aliasing" a two-dimensional data set such as a seismic section.

A Straight Line Model for the Power Spectrum of a Section

Our model of a seismic section will be that it consists of a stochastic collection of straight line segments. Given these, a guess as to the Fourier structure of a stochastic collection of straight line segments can be made. This will prove useful in the analysis of the algorithm which is the subject of this paper.

Suppose we are given N slopes s_l, $l=1,2,3,...,N$. For each slope we are given N_l lines with intercepts t_{jl} and amplitudes a_{jl} , $j=1,2,...,N_l$. The model section corresponding to this collection is

$$\sum_{l=1}^{N}\sum_{j=1}^{N_l} a_{jl}\ \delta(t - t_{jl} - xs_l) \tag{8}$$

which has a two dimensional Fourier transform

$$\sum_{l=1}^{N}\sum_{j=1}^{N_l} a_{jl}\ e^{-i\omega t_{jl}}\ \delta(k + \omega s_l) \tag{9}$$

and power spectrum

$$\sum_{l=1}^{N} \delta(k + \omega s_l) \sum_{j=1}^{N_l}\sum_{k=1}^{N_l} a_{jl}\ a_{kl} \cos\left[\omega\,(t_{kl} - t_{jl})\right] \tag{10}$$

Equation (8) is a superposition of a collection of amplitude modulated line segments. If we assume a uniform distribution of intercepts we can take the expected value of the power spectrum.

$$\sum_{l=1}^{N} \delta(k + \omega s_l) \sum_{j=1}^{N_l} a_{jl}^2 \tag{11}$$

In a stochastic world of line segments the amplitude modulation in equation (11) vanishes. This is a little extreme for data from the real world, so it will instead be assumed that the power spectrum of a section is given by lines with intercepts at the origin (if negative frequencies appear after positive frequencies then there will appear to be two origins) in the FK plane and slow amplitude variations in the radial direction.

Straight lines which are aliased will contribute a straight line segment to the power spectrum of the section. In addition there will be segments at high temporal frequencies which intercept the $\omega = 0$ line at $k = 2m\pi$, where m is an integer other than 0 or 1.

An Iterative Algorithm for Alias Picking and Interpolating in Two Dimensions

Equation (8) implies that a reasonable model for a deconvolved seismic section's power spectrum is that it consists of straight line segments passing through the origin of the FK-plane. Smoothing in radial directions will enhance these segments at the expense of other kinds of energy, including the extra streaks due to aliased straight lines in the x,t-plane.

In practice, the largest departure from the straight line model is due to wavelet color. In practice this is removed before smoothing radially and restored after the iterative interpolation step which will be described next.

The variables in the 2-D problem will have two indices. The first index will be understood to index temporal frequency. There will be understood to be $N\omega + 1$ such temporal frequency points. Similarly, the second index will index the spatial frequency k_x from 0 to Nx. With these conventions we define:

$$\begin{aligned}
&F_{ij} \rightarrow \textit{input transform} && i = 0,1,\cdots,N\omega, \quad j = 0,1,\cdots,Nx \\
&G_{ij} \rightarrow \textit{output transform} && i = 0,1,\cdots,N\omega, \quad j = 0,1,\cdots,2\,Nx \\
&|H_{ij}| \rightarrow \textit{desired spectrum} && i = 0,1,\cdots,N\omega, \quad j = 0,1,\cdots,2\,Nx
\end{aligned} \tag{12}$$

The output section should have a power spectrum which is as close as possible to the desired power spectrum obtained by smoothing the input in radial directions. It should also honor the input data. The way to honor the input data in situations where there is only two-fold aliasing is to use the information contained in equations (1) and (2). One way to do this is to define a norm and minimize it. This norm will have some weights in it, so we introduce

$$W_{ij} \rightarrow weights \qquad i = 0,1, \cdots ,N\omega, \quad j = 0,1, \cdots ,2\,Nx$$

It now makes sense to construct and minimize a norm N, which we define by

$$N = \sum_{j=0}^{Nx-1} \sum_{i=0}^{N\omega} \left\{ W_{ij} \left(|G_{ij}|^2 - |H_{ij}|^2 \right)^2 + W_{ij+Nx} \left(|G_{ij+Nx}|^2 - |H_{ij+Nx}|^2 \right)^2 \right\}$$

or, preserving the input data explicitly,

$$N = \sum_{j=0}^{Nx-1} \sum_{i=0}^{N\omega} \left\{ W_{ij} \left(|G_{ij}|^2 - |H_{ij}|^2 \right)^2 + W_{ij+Nx} \left(|F_{ij} - G_{ij}|^2 - |H_{ij+Nx}|^2 \right)^2 \right\} \tag{13}$$

Equation (13) has some nice properties. It is non-negative and is zero only when $|G_{ij}|$ is equal to $|H_{ij}|$, the desired power spectrum. Estimates of the phase spectrum of the input need not be made. If N is minimized by a gradient technique then the resulting algorithm has a cost which varies roughly linearly with the number of unknowns. In fact, the iterations generally take less time then the smoothing needed to construct $|H_{ij}|$.

To construct a gradient algorithm we take the derivative of N with respect to g_{ij} and the second derivative of N with respect to g_{ij}. This is equivalent to diagonalizing the Hessian matrix and therefore some mistakes in choice of descent direction, but is much cheaper and easier then carrying along the Hessian or its inverse. Differentiating and ignoring the delta functions which result from differentiation of the absolute value function,

$$\frac{\partial N}{\partial G_{ij}} = 4\,W_{ij} \left(|G_{ij}|^2 - |H_{ij}|^2 \right) G_{ij}^{*} - 4\,W_{ij+Nx} \left(|F_{ij} - G_{ij}|^2 - |H_{ij+Nx}|^2 \right) \left(F_{ij}^{*} - G_{ij}^{*} \right) \tag{14}$$

$$\frac{\partial^2 N}{\partial G_{ij}^2} = 8\,W_{ij}\,G_{ij}^{*2} + 8\,W_{ij+Nx} \left(F_{ij}^{*} - G_{ij}^{*} \right)^2 \tag{15}$$

If at a given iteration we have an estimate of G_{ij}, then we can update this estimate by adding an increment δG_{ij} to it.

$$\delta G_{ij} = - \left(\frac{\partial^2 N}{\partial G_{ij}^2} \right)^{-1} \left(\frac{\partial N}{\partial G_{ij}} \right) \tag{16}$$

Preprocessing

To get the algorithm sketched above to work even poorly it is necessary to obtain the best possible power spectral estimates. The input should therefore be preprocessed to remove as many truncations and boundaries as possible. It should also be made to look close to Gaussian.

The first step taken was to remove spherical spreading from the data by multiplying by time. This made the input more stationarity with minimal cost. The data we chose to work with had a sea floor on it which introduces a hard edge into the problem. This was removed next by shifting traces vertically. The shifts introduced a truncation at the bottom of the section. To soften this edge, some extrapolation was done off the edge of the data.

Radix 2 Fourier transform algorithms require the input to be a power of 2 long. It is usually necessary to pad the data to be transformed with zeros. This introduces another hard edge, which was again softened by extrapolation into the padding.

Finally, least squares algorithms yield an optimum estimate when their output is a Gaussian random process. The input data was therefore mapped as close to Gaussian as possibly could be achieved with a power law map of the form $x \rightarrow |x|^{\alpha} sgn(x)$. The parameter α, chosen by manipulation of an empirically obtained cumulative distribution function, is usually close to 0.4 or 0.5. The importance of "Gaussianizing" to alias picking and interpolating cannot be overestimated.

Power Spectrum Weights

The norm in this scheme has some parameters which must be user supplied. These are the weights, W_{ij}, in equation (13). In the example W was chosen as a ramp in the kx direction. More weight was placed on high kx values since it was held to be desirable to force the power spectrum there to zero. Presumably, a similar emphasis at high dips and high temporal frequencies is desirable. The optimum weight may well be the inverse of the expected power spectrum of the output.

Theoretical Extensions

One restriction on this process is that it assumes two-fold aliasing (or no aliasing). Higher aliases can be accounted for by introducing additional unknowns into the norm of equation (13). The norm for this problem might look like

$$N = \sum_{j=0}^{Nx-1}\sum_{i=0}^{N\omega}\left(|G_{ij}|^2 - |H_{ij}|^2\right)^2 + \left(|G_{ij+Nx}|^2 - |H_{ij+Nx}|^2\right)^2 +$$

$$\sum_{j=0}^{Nx-1}\sum_{i=0}^{N\omega}\left(|F_{ij} - G_{ij} - G_{ij+Nx}|^2 - |H_{ij+2Nx}|^2\right)^2$$

for three-fold aliasing.

Another extension which can be made is to change the norm so that it pays more or less attention to peaks in the power spectrum. This can be done by changing some or all of the exponents in the norm. For example, the norm might be chosen to be

$$N = \sum_{j=0}^{Nx-1}\sum_{i=0}^{N\omega}\left\{W_{ij}\left(|G_{ij}|^\beta - |H_{ij}|^\beta\right)^\gamma + W_{ij+Nx}\left(|F_{ij} - G_{ij}|^\beta - |H_{ij+Nx}|^\beta\right)^\gamma\right\}$$

Interpolation Example

The algorithm sketched above was tried on a data set with some crossing events to observe its behavior. Aliasing was not introduced at this time to make further program developments and analysis easier. With this sort of data set a classical interpolator should be close to optimum. The power spectrum at high kx's should be near or equal to zero. It proved possible, with a reasonable weighting function, to get this kind of interpolatory behavior.

The input data is a near offset section corrected for spherical spreading. The effect of the spherical spreading correction is to make the data more stationary. Subsequent estimates of the power spectrum will be better if this is so. Most of the energy is in the sea-floor primary and its first two multiples. There are, in fact, some additional primaries which interfere with the multiples. This will be more evident after the data has been processed to look Gaussian. The input is in Figure 1.

The next step in the preprocess sequence is to shift the sea-floor up to $t = 0$. This was done by picking the first significant peak of the envelope of the first trace. The remaining traces were picked by cross-correlating with this trace. Each trace's shift was saved so that the shift step could be undone at the end.

This shifting and the process of padding to a power of two created a blank at the bottom of the data set. The hard edge was softened with an extrapolation into the padding. The extrapolation was done with a prediction error filter generated by Burg's algorithm.

It was also necessary to pad the sides to a power of two, creating another bunch of anomolous zeros and another hard edge. Extrapolation in this region was done by applying an exponential decay to the first and last traces of the input data. Periodicity was assumed since a Fourier transform was planned a few steps hence.

The result of the shifting, padding, and extrapolating is displayed in Figure 2. It is not a desirable data base for interpolation. It would be nice to make the entire data set look Gaussian. As a step in this direction, the cumulative distribution function for one trace was estimated by sorting into bins. The appropriate power law with which to map the distribution function into the error function was found by a least-squares technique. For this particular piece of data the mapping was of the form $x \rightarrow |x|^{0.4296} sgn x$. The result is in Figure 3. After mapping, crossing events in the x,t-plane stand out much more clearly and the overall appearance is much more "even".

With the preprocessing done, the two dimensional FFT of the data was found. The power spectrum in Figure 4 was constructed from the FFT by taking the square root of the absolute value squared of the transform. The temporal frequency axis is vertical with increasing frequency downwards. Only positive temporal frequencies appear since the input data was real. Spatial frequency is indexed by the horizontal axis. Nyquist in this case is in the middle of the data set.

One of the most obvious features of Figure 4 is the prominent horizontal streaking. This was interpreted as a wavelet effect. To get rid of it, the data was averaged in the k_x direction. The data was then divided by this ω-dependent average and plotted in Figure 5. The model of the output power spectrum was created by smoothing the deconvolved output in radial directions. The length of the averaging window was roughly 18 traces long. Changing the length did not change the output very much. Considerable care was devoted to taking the periodicity of the power spectrum into account during the smoothing process. After smoothing, the wavelet was put back into the problem by multiplication. The output, in Figure 6, looks like a streaky version of Figure 5.

Minimizing the norm in equation (13) yields the power spectrum of the output, interpolated section which is plotted in Figure 7. The most prominent artifact of this power spectrum is the peak near Nyquist values of k_x at low values of ω. A better choice of weights would have attenuated this peak considerably.

An inverse FFT, and the undoing of the sea-floor shifts brings us to Figure 8. This represents the output of the entire process except for the undoing of the power law transformation that makes the data more Gaussian. The output was left in a roughly Gaussian state because more events were detectable there. Figure 8 should therefore be compared with Figure 3 (as opposed to Figures 1 or 2).

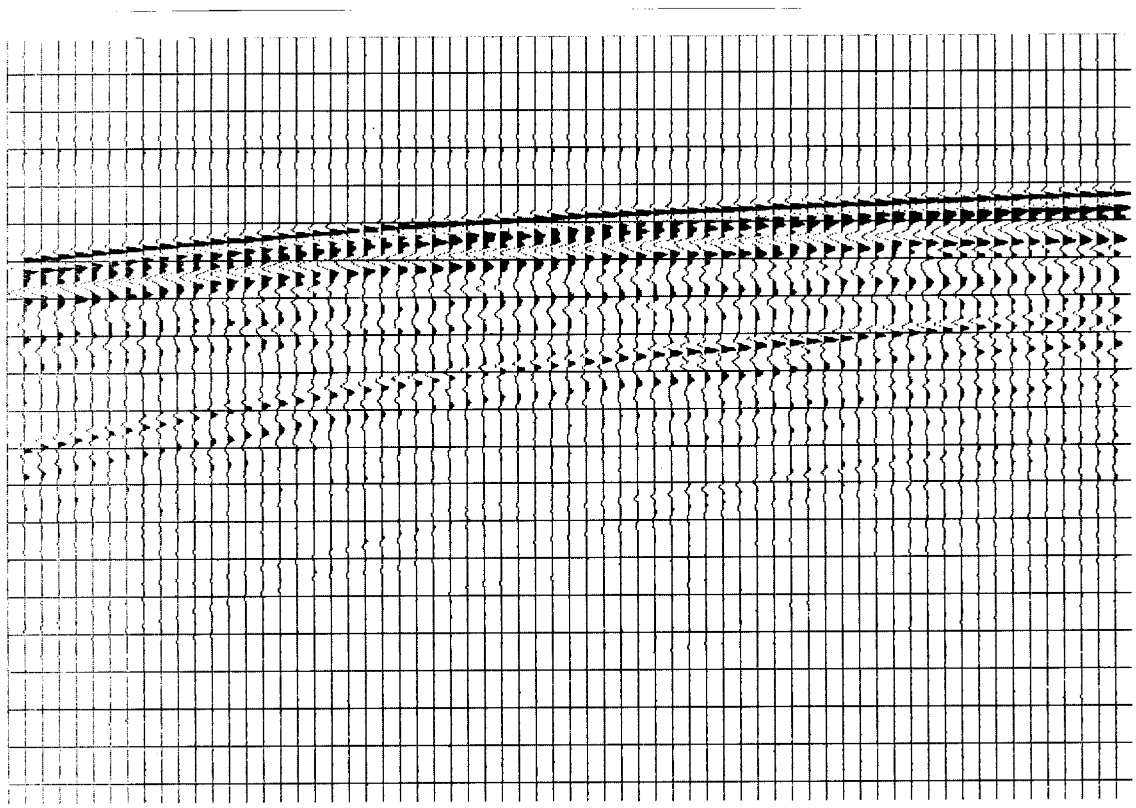

FIG. 1. The input to the interpolation procedure contains a number of interfering events. Spherical spreading correction has been applied.

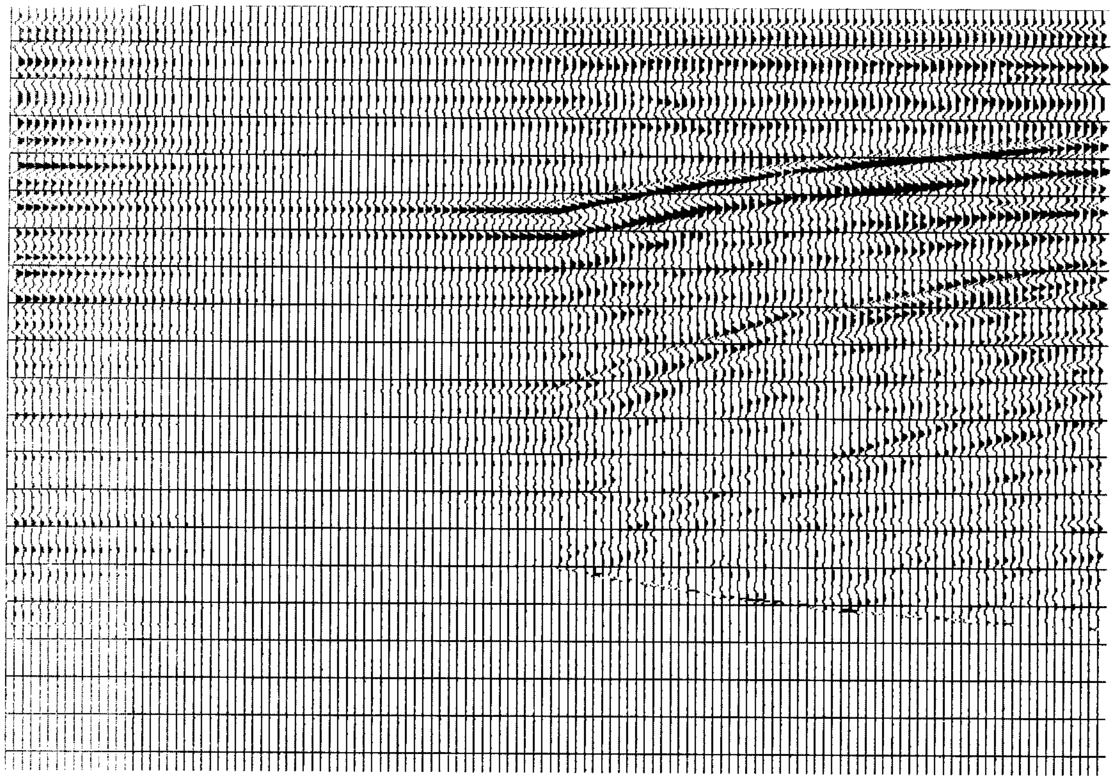

FIG. 2. After shifting, padding, and extrapolating into the padding the input data is just about ready for Fourier transforming.

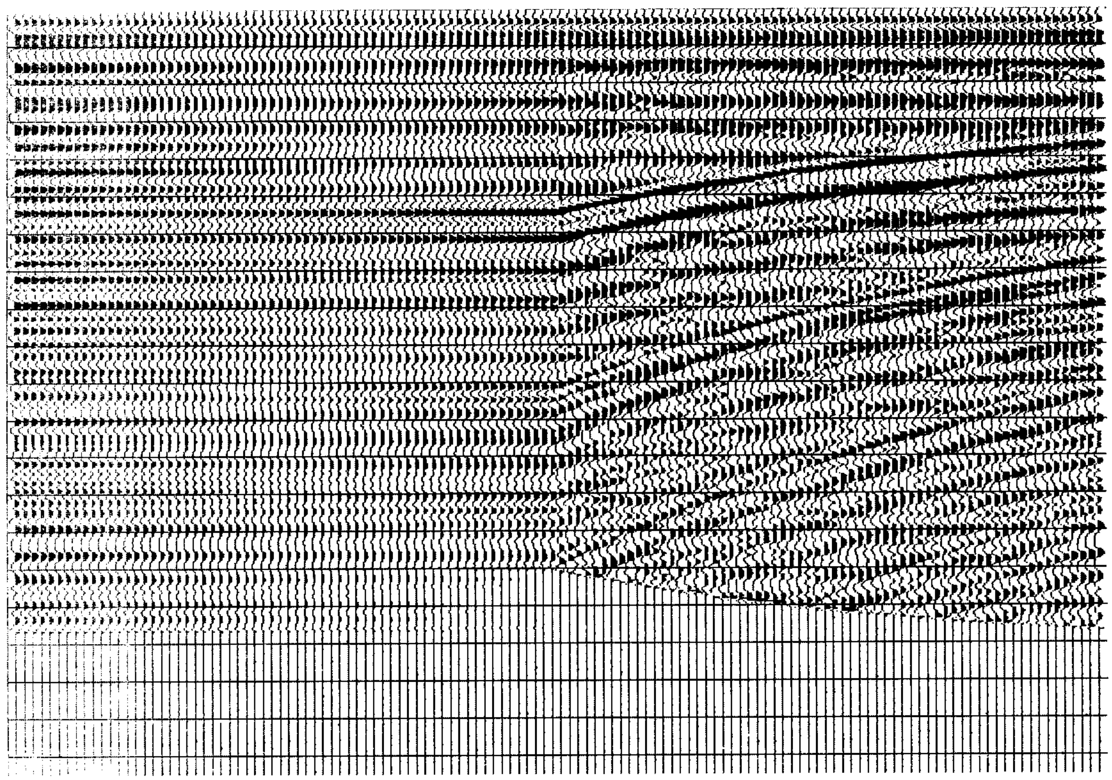

FIG. 3. The padded data set is mapped according to a power law to make it look more Gaussian.

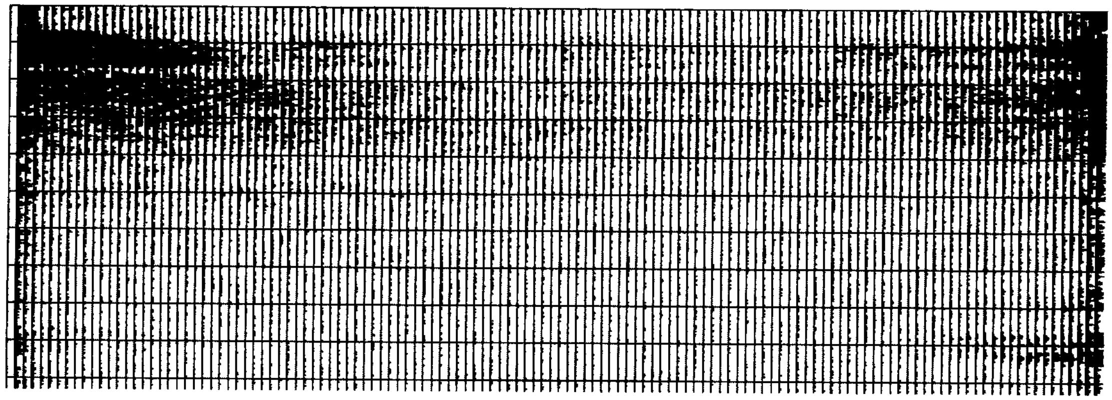

FIG. 4. The power spectrum of the section has prominent horizontal streaks due to the source wavelet and filters applied during acquisition. At this point it does not look too streaky in the radial direction.

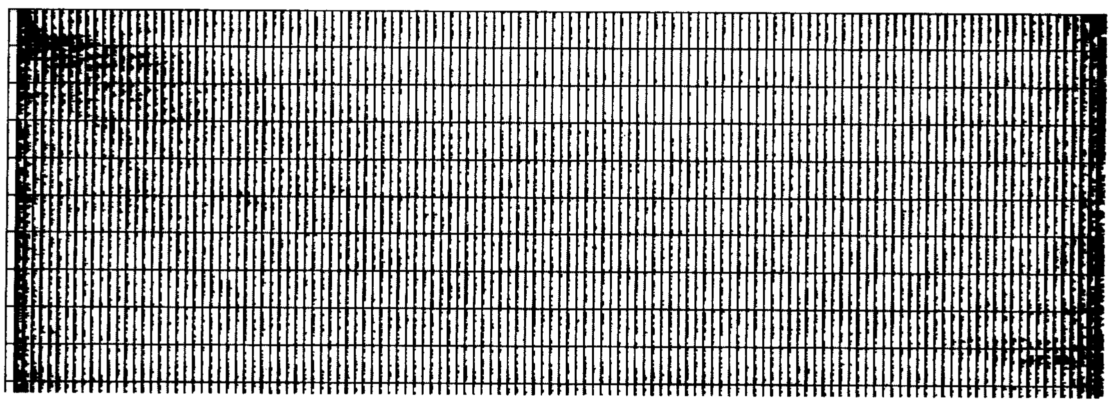

FIG. 5. After removing source wavelet effects, the power spectrum begins to show a little radial streakiness.

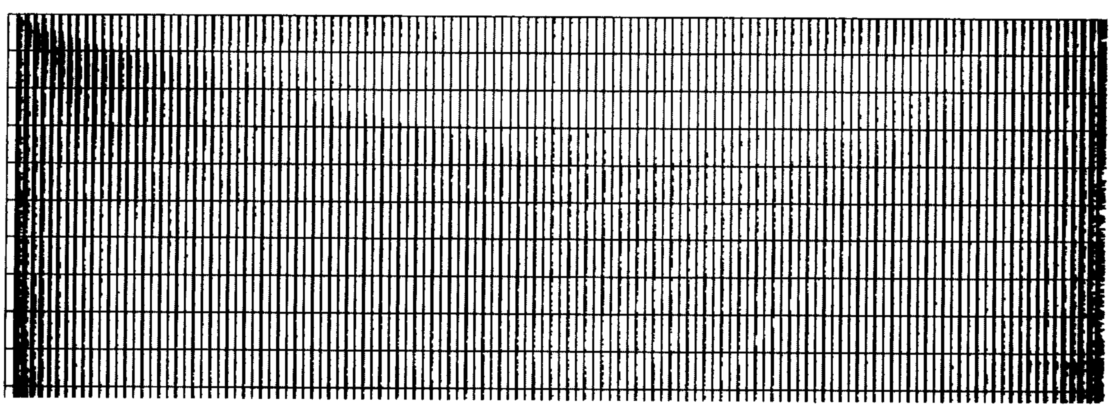

FIG. 6. Radially smoothing the deconvolved power spectrum enhances the streakiness.

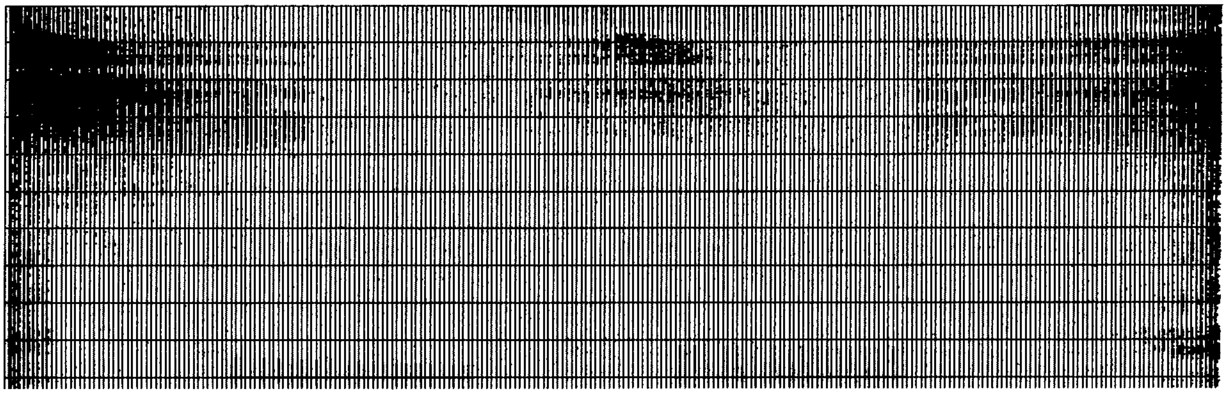

FIG. 7. The power spectrum of the output is close to what was expected. The region of high k_x's, in the middle of the plot is largely filled with small amplitudes. Exceptions which are near Nyquist in k_x are probably artifacts.

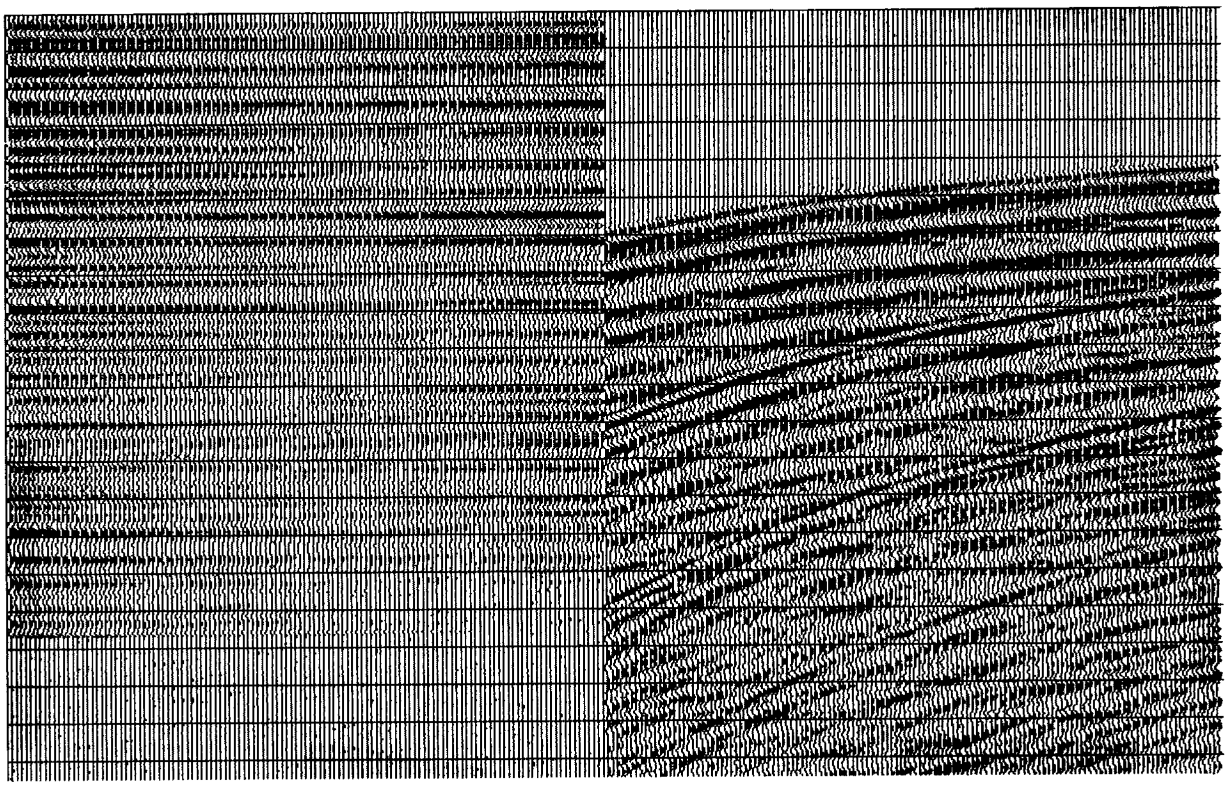

FIG. 8. The output of the process looks interpolated even where events crossed in the input. This plot is comparable to that in Figure 3.

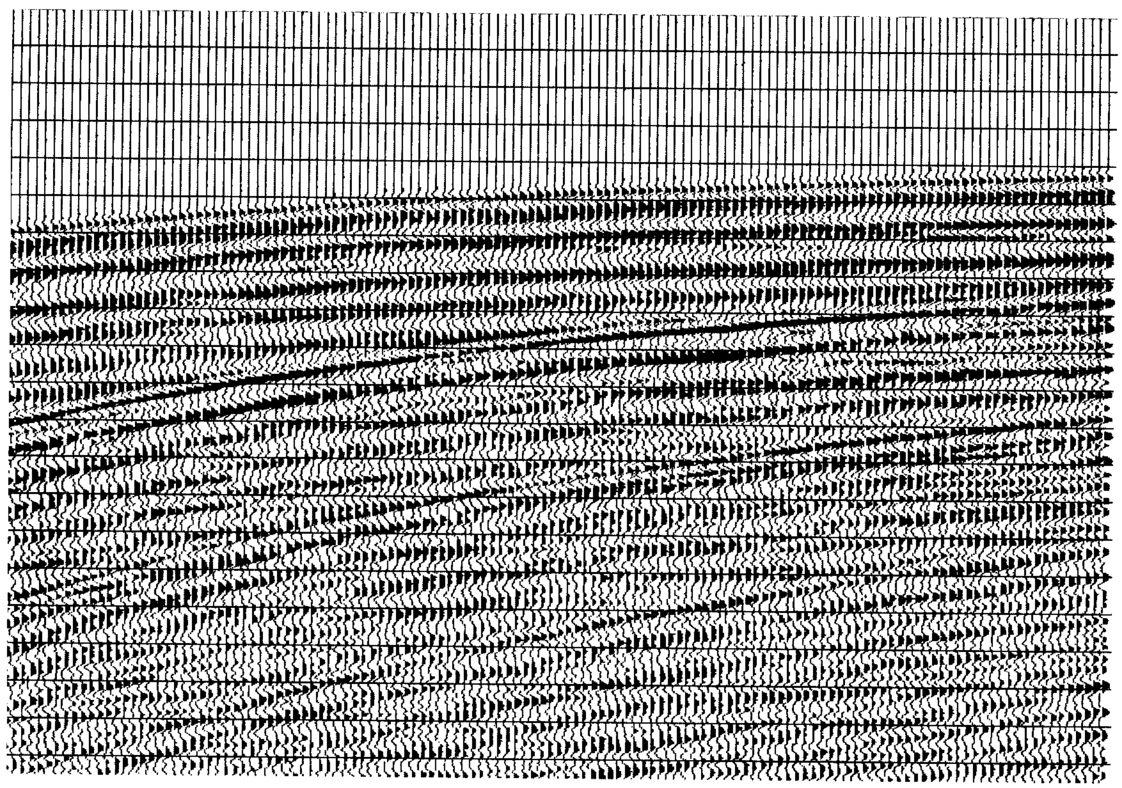

FIG. 9. This is the same data as plotted in Figure 8. The scale is the same as that in Figure 3. The padding has been chopped off.