GEOPHYSICS, VOL. 66, NO. 2 (MARCH-APRIL 2001); P. 667–678, 17 FIGS.

DOI:10.1190/1.1444957

Parabolic Radon transform, sampling and efficiency

M. A. Schonewille* and A. J. W. Duijndam‡

ABSTRACT

A good choice of the sampling in the transform domain is essential for a successful application of the parabolic Radon transform. The parabolic Radon transform is computed for each temporal frequency and is essentially equivalent to the nonuniform Fourier transform. This leads to new and useful insights in the parabolic Radon transform.

Using nonuniform Fourier theory, we derive a minimum sampling interval for the curvature parameter and a maximum curvature range for which stability is guaranteed for general (irregular) sampling. A significantly smaller sampling interval requires stabilization. If diagonal stabilization is used, no gain in resolution is obtained.

In contrast to conventional implementations, the curvature sampling interval is proposed to be inversely proportional to the temporal frequency. This results in improved quality of the transform and yields significant savings in computation time.

INTRODUCTION

The least-squares parabolic Radon transform is a popular transform for suppressing multiples and reconstructing missing traces. When applying the parabolic Radon transform, the sampling of the transform domain, which is defined by the sampling interval in the transform domain (Δq) and the curvature range (Q), must be chosen. A good choice of these parameters is essential for the performance of the parabolic Radon transform and is the principal subject of this paper. The aspects considered are aliasing, stability of the inversion, resolution, and efficiency.

Aliasing in the parabolic Radon transform is discussed by Hugonnet and Canadas (1995) and by Marfurt et al. (1996). The resolution for both the parabolic and linear Radon transform is discussed by Gulunay (1990).

Efficiency can be improved by using the Toeplitz structure of the operators involved and, since the operators are independent of the data, by precomputing the operator and applying it to a number of data gathers with the same geometry. Kelamis and Chiburis (1992) propose a partial stack to regularize the data, which makes the use of precomputed operators possible.

We (1) give a short review of the parabolic Radon transform and the nonuniform Fourier transform, (2) show that for each temporal frequency the transforms have an equivalent structure, (3) use nonuniform Fourier theory to derive the maximum Q and minimum Δq for which stability is guaranteed, and (4) investigate the effects of using a smaller Δq or larger Q while using diagonal stabilization. We propose using a Δq that is inversely proportional to the temporal frequency (consequently, the complete transform becomes independent of frequency). The frequency-dependent Δq leads to more efficient algorithms and provides a natural way to handle the lower temporal frequencies better. Finally, we give a synthetic example in which the new algorithm is compared with the conventional parabolic Radon transform.

PARABOLIC RADON TRANSFORM

This section is a short overview of the parabolic Radon transform. A more comprehensive overview can be found in Zhou and Greenhalgh (1994) and in Dunne and Beresford (1995).

Forward transform

The discrete parabolic Radon transform was introduced by Hampson (1986). The direct forward transform is defined as

$$m(q, \tau) = \sum_{n=1}^{N} d\left(x_n, t = \tau + q x_n^2\right), \qquad (1)$$

where $d(x_n, t)$ is the data in the offset(x)-time (t) domain, N is the number of traces, and $m(q, \tau)$ is the data in the parabolic Radon domain. In the parabolic Radon domain the data are a function of the curvature q and the zero offset intercept time τ.

Manuscript received by the Editor December 29, 1998; revised manuscript received June 30, 2000.
*Formerly Delft University of Technology, Faculty of Applied Earth Sciences, Delft, The Netherlands; presently PGS Seres, PGS Court, Halfway Green, Walton-on-Thames, Surrey KT12 1RS, United Kingdom. E-mail: michel.schonewille@london.pgs.com.
‡Formerly Delft University of Technology, Faculty of Applied Earth Sciences, 2600 GA Delft, The Netherlands; presently Philips Medical Systems, PMG Magnetic Resonance, P.O. Box 10.000, 5680 DA Best, The Netherlands. E-mail: Adri.Duijndam@philips.com.

After a temporal Fourier transformation, the parabolic Radon transform can be calculated for each temporal frequency component ω:

$$M(q,\omega)=\sum_{n=1}^{N} D(x_n,\omega)\exp\left(j\omega q x_n^2\right), \tag{2}$$

where $M(q,\omega)$ and $D(x_n,\omega)$ are the temporal Fourier transforms of $m(q,\tau)$ and $d(x_n,t)$, respectively.

Inverse transform

An approximate inverse of relation (2) is

$$D(x_n,\omega)=\sum_{i=1}^{N_q} M(q_i,\omega)\exp\left(-j\omega q_i x_n^2\right), \tag{3}$$

where N_q is the number of q values. Using matrix notation, equation (3) can be written as

$$\mathbf{d}=\mathbf{L}\mathbf{m}, \tag{4}$$

with

$$m_i = M(q_i,\omega), \qquad i=1,\ldots,N_q, \tag{5}$$

$$L_{ni}=\exp\left(-j\omega q_i x_n^2\right), \tag{6}$$

$$d_n = D(x_n,\omega), \qquad n=1,\ldots,N. \tag{7}$$

Least-squares formulation

Instead of using the transform pair given by equations (2) and (3), it is common practice to replace the forward transform with a least-squares transform, giving better reconstruction and better resolution in the Radon domain which allows better signal and noise separation.

The least-squares forward transform is derived by using the inverse transform [equations (3) and (4)] as a forward model in a least-squares formulation:

$$\hat{\mathbf{m}}=(\mathbf{L}^H\mathbf{L})^{-1}\mathbf{L}^H\mathbf{d}, \tag{8}$$

in which $\hat{\mathbf{m}}$ is the estimated data in the Radon domain.

Sampling in the transform domain

Several authors have studied sampling in the transform domain [see Hugonnet and Canadas (1995) and references cited therein]. For regular sampling in x, $x_n=x_{\min}+n\Delta x$ and $n=[0,\ldots,N-1]$, the well-known formulas for the sampling interval Δq and the maximum q-range that can be estimated are, respectively,

$$\Delta q < \frac{2\pi}{\omega\left(x_{\max}^2-x_{\min}^2\right)}, \tag{9}$$

where $x_{\max}=x_{\min}+(N-1)\Delta x$ is the maximum offset, and

$$Q_{lsip}=\frac{\pi}{\omega x_{\max}\Delta x}, \tag{10}$$

where Q_{lsip} is the curvature range for which local summation in phase occurs (for the largest offsets; see Hugonnet and Canadas, 1995). In conventional parabolic Radon transform implementations Δq is chosen independent of the frequency by setting $\omega=\omega_{\max}$ in equations (9) and (10).

If the data are not aliased, then the q-range can be extended by applying operator dealiasing (Cary, 1998). Although this does not allow for plain parabolic events (wavelet independent of offset) with a curvature range larger than Q_{lsip} (it can be derived that in this case the data will be spatially aliased), it can give a significant improvement in case of an offset-dependent frequency content (e.g., from NMO stretching) and in case of smearing of energy in the Radon domain from amplitude versus offset (AVO) effects, for example.

NONUNIFORM FOURIER TRANSFORM

The parabolic Radon transform is very closely related to the nonuniform Fourier transform. Consider a set of nonuniformly spaced, distinct, and ordered sample positions ($x_1<x_2<\cdots<x_N$). For a band-limited signal we can use the discrete inverse Fourier transform as a model for the data $P(x_n,\omega)$:

$$P(x_n,\omega)=\frac{\Delta k}{2\pi}\sum_{m=0}^{M-1}\tilde{P}(m\Delta k,\omega)\exp(-jm\Delta k x_n), \qquad n=1,\ldots,N, \tag{11}$$

where $\tilde{P}(m\Delta k,\omega)$ are the unknown Fourier coefficients. By sampling the Fourier domain with sampling interval Δk, we implicitly consider the data set to be one period of a periodic signal. The periodicity interval is given by

$$X=2\pi/\Delta k. \tag{12}$$

Equation (11) can be written in matrix notation as

$$\mathbf{p}=\mathbf{A}\tilde{\mathbf{p}}, \tag{13}$$

with

$$\tilde{p}_m = P(m\Delta k,\omega), \qquad m=0,\ldots,M-1, \tag{14}$$

$$A_{nm}=\frac{\Delta k}{2\pi}\exp(-jm\Delta k x_n), \tag{15}$$

and

$$p_n = P(x_n,\omega), \qquad n=1,\ldots,N. \tag{16}$$

The nonuniform forward Fourier transform (Feichtinger et al., 1995; Duijndam et al., 1999) is now defined as the weighted least-squares inverse of equation (13):

$$\hat{\tilde{\mathbf{p}}}=(\mathbf{A}^H\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{W}\mathbf{p}, \tag{17}$$

where $\mathbf{W}$ is a diagonal matrix with diagonal elements

$$W_{ii}=\frac{x_{i+1}-x_{i-1}}{2}. \tag{18}$$

For band-limited signals this transform is exact, apart from spatial aliasing introduced by the sampling of the transform domain. The matrix $\mathbf{A}^H\mathbf{W}\mathbf{A}$ is theoretically invertible as long as the number of distinct samples N is larger than or equal to the number of Fourier coefficients M, without any restriction to the actual sample positions (see Bagchi and Mitra, 1996). However, in the case of large gaps in the sampling positions, the conditioning of $\mathbf{A}^H\mathbf{W}\mathbf{A}$ is very bad; in practice, a good reconstruction of the signal is impossible in the gaps (see, e.g., Duijndam et al., 1999).

For a proper Fourier transform, or reconstruction, of the signal, the choices of the sampling interval in the Fourier domain

Δk and the number of Fourier coefficients M is of great importance. Our objective is to achieve an accurate reconstruction of the signal. This means that we should have a stable inversion and a good data fit. If the data are not fitted well, the forward model does not appropriately describe the data. It can be expected that the data fit becomes better if the k-range in the model is increased. The objective then implies that we would like to use the largest k-range that can be estimated in a stable way while avoiding spatial aliasing.

The choices of Δk and M are ultimately related to the gaps between the sample positions. Besides the gaps between the actual samples, we also introduce a gap Δx_i through the choice of Δk between the last sample in the set, x_N, and the first sample in the next period, $x_1 + X$:

$$\Delta x_i = x_1 + X - x_N = \frac{2\pi}{\Delta k} - X_a, \tag{19}$$

where $X_a = x_N - x_1$ is the actual aperture of the sampling set. The largest gap in the sampling set is

$$\delta = \max(\Delta x_i, \Delta x_a), \tag{20}$$

where Δx_a is the largest gap between the actual sample positions. Gröchenig (1993) proved that the condition number κ of the matrix $\mathbf{A}^H\mathbf{W}\mathbf{A}$ for general irregular sampling satisfies

$$\kappa \le \left(\frac{\dfrac{2\pi}{(M-1)\Delta k} + \delta}{\dfrac{2\pi}{(M-1)\Delta k} - \delta} \right)^2 \quad \text{for } \delta < \frac{2\pi}{(M-1)\Delta k}. \tag{21}$$

The smaller the condition number, the more stable the inversion. Note that equation (21) represents an upper bound for κ; often, the conditioning is better than this upper limit suggests. We show, however, that the conditioning of the least-squares parabolic Radon transform for regular spatial sampling is fairly well described by equation (21). Therefore, let's study the implications of equation (21) further.

The general behavior of the upper bound of the condition number as a function of M and Δk is illustrated in Figure 1 for the case $\Delta x_a = 0.04$ and $X_a = 1$. For fixed M, the upper bound of the condition number increases if Δk decreases and goes to infinity at the white line, where $\delta = 2\pi/[(M-1)\Delta k]$. Note that the upper bound is defined only to the right of the white line, and it is clipped to 1000 for display purposes. For

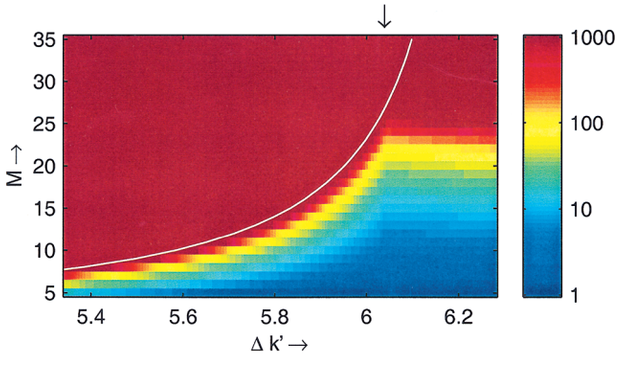

FIG. 1. Upper bound of the condition number for Fourier reconstruction as a function of the sampling interval in the transform domain, Δk, and the number of estimated k-values, M.

sufficiently large M the inversion always becomes unstable, which is obvious.

With respect to Δk we can distinguish two regions:

1) The region for which the gap that is introduced is larger than the maximum actual gap ($\Delta x_i > \Delta x_a$), which is the case for

$$0 < \Delta k < \frac{2\pi}{X_a + \Delta x_a}, \tag{22}$$

and

2) its complement, $\Delta x_i \le \Delta x_a$, for

$$\frac{2\pi}{X_a + \Delta x_a} \le \Delta k < \frac{2\pi}{X_a}. \tag{23}$$

The upper bound $\Delta k < 2\pi/X_a$ stems from the fact that beyond that range the periodicity interval $X = 2\pi/\Delta k$ is smaller than the actual aperture X_a and serious edge effects are introduced.

From equation (21) we can derive that the maximum number of Fourier coefficients that can be estimated (and the largest k-range, to a good approximation) is obtained at the boundary between these two regions,

$$\Delta k = \frac{2\pi}{X_a + \Delta x_a}, \tag{24}$$

for which $\Delta x_i = \Delta x_a$. In Figure 1 this Δk is marked by the arrow at the top. In the next section we will show that for the parabolic Radon transform the Δq related to Δk also provides a good data fit.

With equations (24) and (21), a theoretical upper bound for M is

$$M < X_a/\Delta x_a + 2. \tag{25}$$

Relations (24) and (25) specify the sampling in the transform domain, based on the upper bound for the condition number, equation (21). For the special case of regular sampling ($x_n = n\Delta x$), these results correspond to the standard discrete Fourier transform choices:

$$\Delta k = \frac{2\pi}{N\Delta x} \tag{26}$$

and

$$M < N + 1. \tag{27}$$

If, from prior information, we know that we can reduce the k-range, then we can decrease Δk and M.

PARABOLIC RADON TRANSFORM AND THE NONUNIFORM FOURIER TRANSFORM

In the parabolic Radon transform we estimate, for each ω, $M(q_i, \omega)$ in a least-squares sense from equation (3), repeated here for convenience:

$$D(x_n, \omega) = \sum_{i=1}^{N_q} M(q_i, \omega) \exp\left(-j\omega q_i x_n^2\right). \tag{28}$$

Using the variable transformations $y_n = x_n^2$ and $k_i' = \omega q_i$, we can write this formula as

$$D_n(\omega) = \sum_{i=1}^{N_q} M(k_i', \omega) \exp(-jk_i' y_n). \tag{29}$$

Comparing equation (29) to equation (11), we see that, for a single ω, the parabolic Radon transform is essentially identical to the nonuniform Fourier transform. This implies that the results concerning the nonuniform Fourier transform can be applied to the parabolic Radon transform. The variable transformation $y_n = x_n^2$ can be interpreted as a quadratic stretching of the x-axis, which turns parabolas into straight lines (see Figure 2).

From the variable transformation $k_i' = \omega q_i$ it follows that a constant Δq as used in the conventional parabolic Radon transform yields a sampling in k' with smaller sampling intervals for lower temporal frequencies (Figure 3). As a consequence, the periodicity in the spatial domain becomes larger and larger for lower frequencies (and a larger gap is introduced).

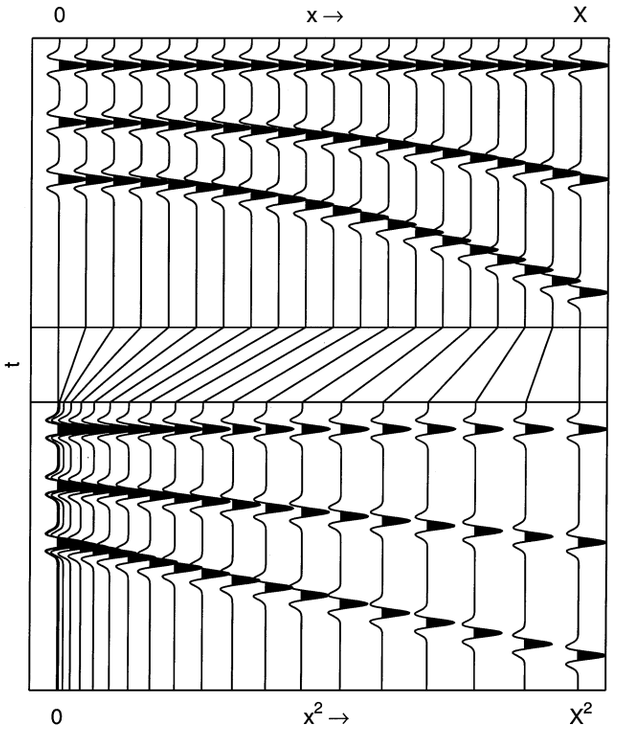

FIG. 2. Stretching of the x-axis. A nonuniform Fourier transform of the stretched data is strongly related to the parabolic Radon transform.

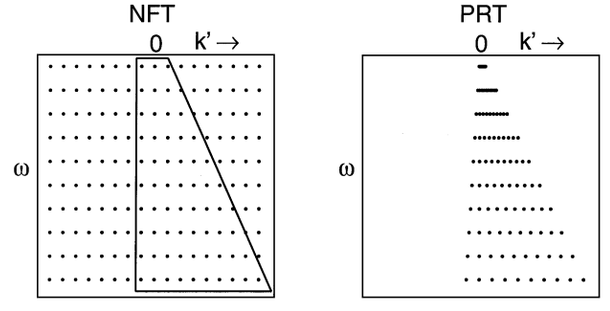

FIG. 3. Sampling in the transform domain: constant $\Delta k'$ for the nonuniform Fourier transform versus constant Δq for the conventional parabolic Radon transform.

Using $k_i' = \omega q_i$ and equation (24), we obtain

$$\Delta q = \frac{2\pi}{\omega(Y_a + \Delta y_a)} = \frac{2\pi}{\omega\left(x_{\max}^2 - x_{\min}^2 + \Delta x_a^2\right)}, \quad (30)$$

where $\Delta x_a^2 = \max(x_{n+1}^2 - x_n^2)$, $n = [1, \ldots, N-1]$ is the maximum gap in sampling positions after the quadratic stretching of the x-axis. We conclude that Δq should depend on ω, and should not be chosen fixed corresponding to $\omega_{\max}$ as in conventional parabolic Radon transform implementations.

Using equation (25), we find

$$M < Y_a/\Delta y_a + 2 = \frac{x_{\max}^2 - x_{\min}^2}{\Delta x_a^2} + 2. \quad (31)$$

These aliasing/stability conditions are valid for general (irregular) spatial sampling.

Let us now look at the case of regular sampling ($x_n = x_{\min} + n\Delta x$, $n = [0, \ldots, N-1]$). Obviously, after the stretching of the x-axis, the largest gap in the actual sampling is between the two largest offsets:

$$\Delta x_a^2 = x_{\max}^2 - (x_{\max} - \Delta x)^2 = 2x_{\max}\Delta x - (\Delta x)^2, \quad (32)$$

where $x_{\max} = x_{\min} + (N-1)\Delta x$. For Δq we find, using equation (30),

$$\Delta q = \frac{2\pi}{\omega\left(x_{\max}^2 - x_{\min}^2 + 2x_{\max}\Delta x - (\Delta x)^2\right)}. \quad (33)$$

Note that equation (33) satisfies equation (9) but is more precise. For M we find, using equation (31) for this case or regular sampling,

$$M < \frac{x_{\max}^2 - x_{\min}^2}{2x_{\max}\Delta x - (\Delta x)^2} + 2. \quad (34)$$

Because normally $x_{\max}^2 \gg x_{\min}^2$ and $x_{\max} \gg \Delta x$, the maximum M that can be estimated is approximately half of the aperture divided by the sampling interval ($x_{\max}/2\Delta x$).

Equations (33) and (34) specify the sampling in the transform domain for the parabolic Radon transform for regular sampling in x and are in fact derived from the stability criterion as given in equation (21), which is displayed in Figure 1 (for general sampling). For comparison, in Figure 4a the numerically determined condition number of the matrix $\mathbf{A}^H\mathbf{A}$ is plotted for the parabolic Radon transform with regular sampling in x. By choosing Δq not fixed but inversely proportional to ω, we are effectively using a constant $\Delta k'$:

$$\Delta k' = \frac{2\pi}{Y_a + \Delta y_a}, \quad (35)$$

which is also indicated by the arrow in Figure 4a. In this example, $x_{\min} = 0$, $\Delta x = 1/49$, and $N = 50$, from which we have $Y_a = 1$ and $\Delta y_a = 0.04$. Again, to the right of the white line, the condition number as given by equation (21) is bounded. Even though we are now considering the unweighted matrix (because that is most commonly used for the parabolic Radon transform), we can conclude that the general behavior closely follows that of the theoretical upper bound (see Figure 1). The major difference is that the actual condition number increases less strongly with increasing M.

For regular sampling in x, the maximum q-range ($M\Delta q$) that can be estimated in a stable way can be derived directly from equations (33) and (34). Note, however, that normally

$x_{\max}^2 \gg x_{\min}^2$ and $x_{\max} \gg \Delta x$ and therefore equation (33) can be approximated by $2\pi/\omega x_{\max}^2$ and equation (34) by $x_{\max}/2\Delta x$, so that we find

$$M\Delta q \approx \frac{x_{\max}}{2\Delta x}\frac{2\pi}{\omega x_{\max}^2} = \frac{\pi}{\omega x_{\max}\Delta x}, \tag{36}$$

which is equal to the conventional formula [equation (10)]. The curvature range is inversely proportional to $x_{\max}$.

If the curvature range of the data is larger than the curvature range that can be estimated in a stable way, then the data will not be fitted well. From equation (36) it follows that both $x_{\max}$ and ω can be reduced to increase the curvature range that can be estimated in a stable way, which can easily be understood from nonuniform Fourier theory. After the x-squared stretch, the largest gaps occur at the largest offsets. Reducing $x_{\max}$ and/or ω is proposed by Hugonnet and Canadas (1995). Better than simply removing large offsets and/or high frequencies, one could determine for each temporal frequency the maximum offset range that could be used in the least-squares inversion. The stack could be normalized in the frequency domain by the effective fold. Note that decreasing the number of samples to increase the k'-range and curvature range strongly reduces the resolution in the transform domain [see equation (33)].

Not only the estimated curvature range but also the Δq (or $\Delta k'$) affects the data fit. This is demonstrated in Figure 4b. The residual energy (the sum of the squares of the difference between the original data and the data after forward and backward transformation) is given as a function of $\Delta k'$ and M, with $N = 50$ and $Y_a = 1$. For each M, a test data set is calculated

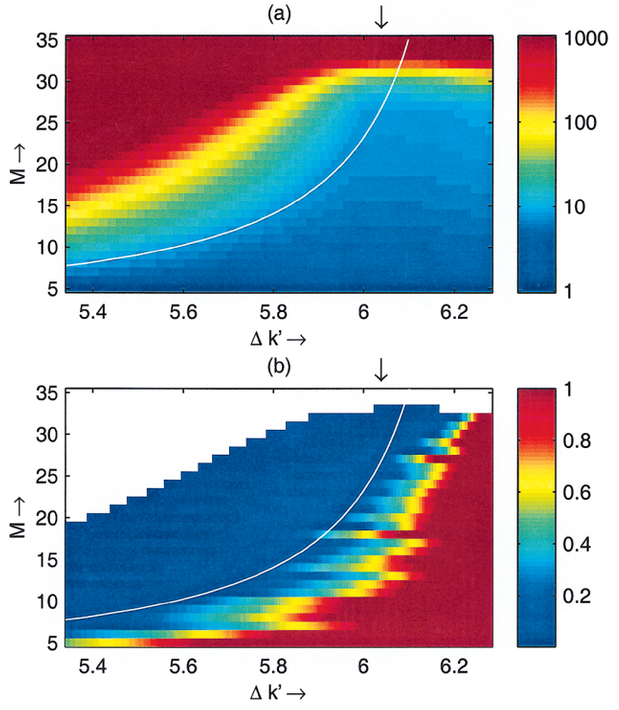

FIG. 4. Parabolic Radon transform: (a) condition number; (b) data fit, both as a function of the sampling interval in the transform domain, $\Delta k'$, and the number of estimated k'-values, M.

with 10 parabolas with random curvatures (within the calculated curvature range) and $\Delta k'$ is varied. The residual energy is scaled so a value of 1 corresponds with the energy of one trace. In the region left of the line, stability is not guaranteed. Remarkably, to the right of this line the data fit starts to decrease. On the line the data fit is good (i.e., the residual energy is much smaller than the energy of one trace), but for the best data fit a somewhat smaller $\Delta k'$ is desired.

DIAGONAL STABILIZATION

It has been shown that both a small $\Delta k'$ and a large k'-range lead to instability. To prevent noise from blowing up, some kind of stabilization must be used. The most commonly used stabilization is diagonal stabilization. In this section the effect of using diagonal stabilization is discussed for both cases of instability.

The diagonally stabilized least-squares transform

$$\hat{\tilde{\mathbf{p}}} = (\mathbf{A}^H\mathbf{A} + \alpha^2\mathbf{I})^{-1}\mathbf{A}^H\mathbf{p} \tag{37}$$

is the least-squares solution of the augmented system of equations

$$\begin{pmatrix}\mathbf{p}\\ \mathbf{0}\end{pmatrix} = \begin{bmatrix}\mathbf{A}\\ \alpha\mathbf{I}\end{bmatrix}\tilde{\mathbf{p}}. \tag{38}$$

The diagonal stabilization implicitly specifies that the signal should be zero everywhere. One may look at this as prior information. The weight of this information is controlled by α. It implies that for spatial locations where the vector equation $\mathbf{p} = \mathbf{A}\tilde{\mathbf{p}}$ does not specify enough information, the diagonal stabilization will push the reconstructed signal to zero.

Instability caused by a large k'-range

In this section, the effect of using diagonal stabilization (with α^2 equal to 1% of the main diagonal value of $\mathbf{A}^H\mathbf{A}$) to handle instability caused by a large k'-range will be demonstrated using a synthetic example. In Figure 5 the absolute amplitude spectrum of a monofrequency data set consisting of one parabolic event with curvature just inside the estimated k'-range is shown for a sufficiently small k'-range. In Figure 6 the parabolic Radon transform spectrum is shown for a k'-range that is too large. The spectrum should have one peak of amplitude one; instead, the energy of the single event spreads over a large part of the spectrum. The interpretation of diagonal stabilization given above helps to explain the spreading of energy. After the quadratic stretching, the sampling interval is larger for larger offsets. If these intervals are too large, then the prior information between the sampling points will dominate the result of the least-squares inversion and the reconstructed signal will be zero. As an illustration, in Figure 7 the result of interpolation to a grid with half the offset sampling interval is shown for the good k'-range and in Figure 8 for the k'-range that is too large. Note that the data fit at the original grid is good in both cases.

Another disadvantage of the spreading of the energy is that specific processing of the multiples in the transform domain will also affect the primaries and vice versa (see also Hugonnet and Canadas, 1995).

It can be concluded that diagonal stabilization is not a proper means to increase the k'-range. Instead, the offset range should be reduced, as described above.

Instability caused by a small $\Delta k'$

When using a small $\Delta k'$, a large gap is introduced in the (implicitly) periodic data between the largest offset and the smallest offset of the next period. After a forward transformation, the data can be reconstructed, in principle, within the complete aperture, defined by $2\pi/\Delta k'$, for the academic case that there is no noise. To illustrate this, a monochromatic single curvature data set is used (see Figure 9). The number of offsets is 40. After the least-squares forward transformation, the data can be reconstructed on a new grid. Figure 10 shows the result for $\Delta k' = 0.75\Delta k'_a$, where $\Delta k'_a = 2\pi/(Y_a + \Delta y_a)$ [see equation (35)], and no diagonal stabilization. Although the condition number is large, using double precision computations we can extrapolate the data well for the offsets of 1000 m and further. In Figure 11 the result is shown in the case of diagonal stabilization. Now the data have not been reconstructed beyond the original aperture. The diagonal stabilization (again 1%) forces the reconstruction to zero in large gaps. In the next section we demonstrate that, as a consequence, the resolution in the transform domain is not higher than for the choice of $\Delta k' = \Delta k'_a$.

RESOLUTION

In this section the equivalence between the nonuniform Fourier transform and the parabolic Radon transform will be used to study the resolution of the parabolic Radon transform. We will investigate whether a smaller $\Delta k'$ can lead to a better resolution.

To find analytical expressions for the response of the parabolic Radon transform, we use the discrete Fourier transform of regularly sampled data. This is possible because band-limited, irregularly sampled data can be reconstructed on a regular grid by

$$\mathbf{p}_r = \mathbf{A}_r(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{p} \tag{39}$$

as long as the signal is sampled sufficiently dense. The matrix $\mathbf{A}_r$ is similar to $\mathbf{A}$ [equation (15)] for regularly spaced x_n:

$$A_{r,nm} = \frac{\Delta k}{2\pi}\exp(-jm\Delta k n\Delta x), \tag{40}$$

where $\Delta x = X/N = 2\pi/N\Delta k$. Note that $\mathbf{A}_r$ represents a standard inverse discrete Fourier transform and therefore

$$\frac{2\pi\Delta x}{\Delta k}\mathbf{A}_r^H\mathbf{A}_r = \mathbf{I}. \tag{41}$$

Combining equations (39) and (41) yields

$$\frac{2\pi\Delta x}{\Delta k}\mathbf{A}_r^H\mathbf{p}_r = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{p}, \tag{42}$$

which shows that the least-squares inverse of the irregularly sampled data is equal to the discrete Fourier transform of the

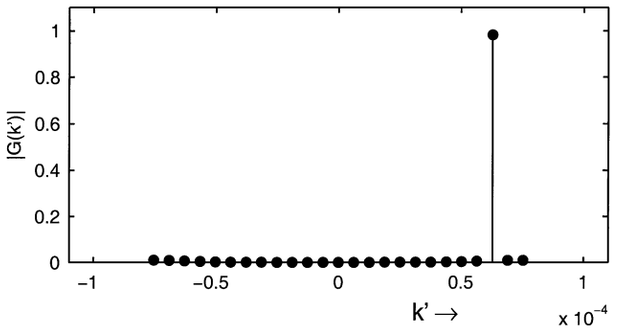

FIG. 5. Absolute amplitude spectrum; k'-range good.

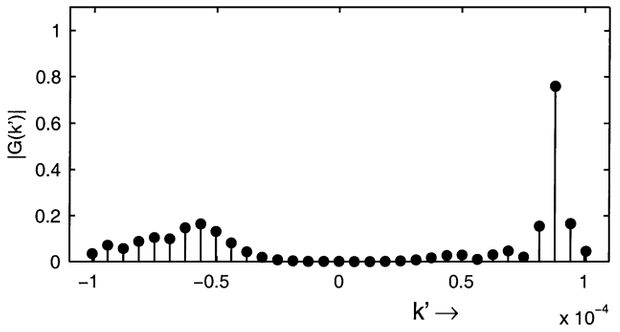

FIG. 6. Absolute amplitude spectrum; k'-range too large.

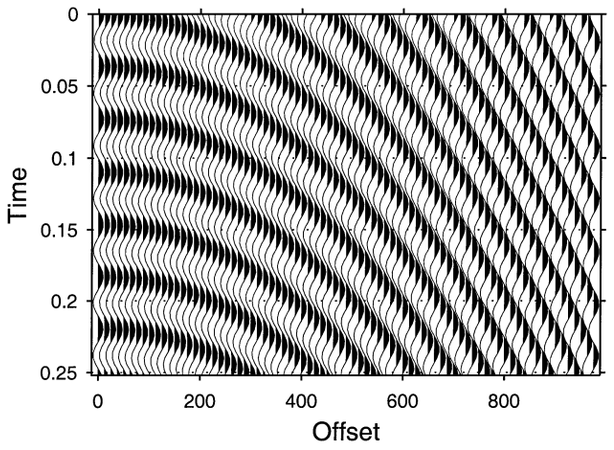

FIG. 7. Interpolated data; k'-range good.

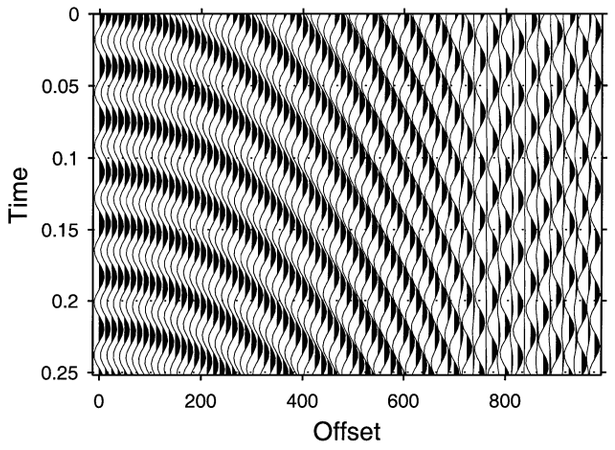

FIG. 8. Interpolated data; k'-range too large.

regularly sampled data. The inverse $(\mathbf{A}^H\mathbf{A})^{-1}$ exists as long as the number of estimated k-values is smaller than or equal to the number of distinct spatial samples.

Resolution in case of no stabilization

For a certain frequency $\omega = \omega_0$, the response of the parabolic Radon transform for a single curvature data set is the same as the response of a single wavenumber data set for the Fourier transform. The response can therefore be calculated by

$$G(k', \omega_0) = \sum_{n=0}^{N-1} P(\omega_0) \exp(j(k' - k'_0)n\Delta y), \qquad (43)$$

where k'_0 is the wavenumber of the data set. If we take the amplitude $P(\omega_0) = 1$ and leave out the ω_0 from $G(k', \omega_0)$, we obtain

$$G(k') = \sum_{n=0}^{N-1} \exp(j(k' - k'_0)n\Delta y). \qquad (44)$$

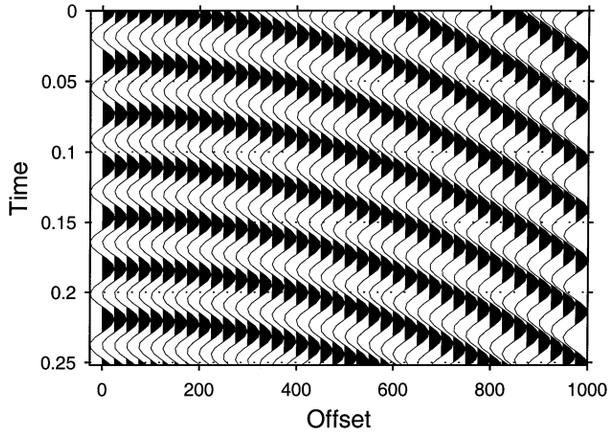

FIG. 9. Single curvature monochromatic data set.

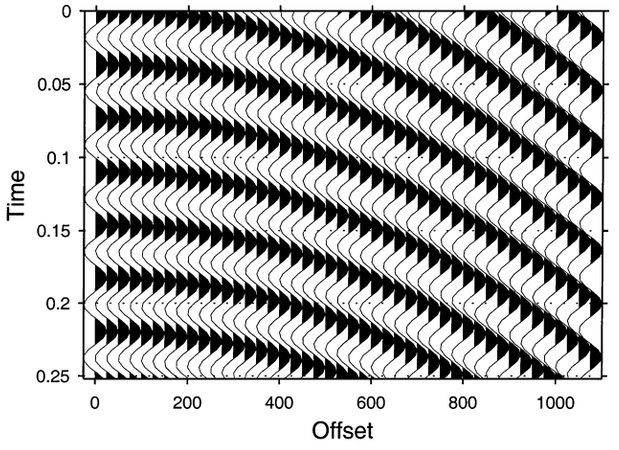

FIG. 10. Reconstructed data with $\Delta k' = 0.75\Delta k'_a$ and no stabilization. The data are extrapolated well.

The absolute value of this function is

$$|G(k')| = \frac{|\sin((k' - k'_0)\Delta y N/2)|}{|\sin((k' - k'_0)\Delta y/2)|}. \qquad (45)$$

The formula $\sin(Nx)/\sin(x)$ is the discrete counterpart of the sinc function $\sin(x)/x$.

For the discrete sinc function two components of equal intensity may be regarded just resolvable if the maximum of one component coincides with the first minimum of the other (see Gulunay, 1990) and the minimum resolvable k'-difference, k'_{diff}, is given by

$$k'_{\text{diff}} = \frac{2\pi}{N\Delta y}. \qquad (46)$$

The resolution is higher if $N\Delta y$ is larger (the aperture is larger). Since for the discrete Fourier transform $\Delta k' = 2\pi/N\Delta y$, this is related to a smaller $\Delta k'$. Therefore, the resolution in principle is also higher if $\Delta k'$ is chosen smaller.

In Figure 12, on the right-hand side, the response is shown without diagonal stabilization and for several choices of $\Delta k'$. For the experiment 40 offsets were used, $\Delta x = 25$. By using a limited k'-range (number of estimated k'-values = 17) and double precision computations, Δk can be chosen as small as $\Delta k' = 0.5\Delta k'_a$ without need for stabilization. It is clear that the resolution improves for smaller $\Delta k'$.

Resolution in case of diagonal stabilization

If diagonal stabilization must be used, then the result is different (see Figure 12, left-hand side). The stabilization factor is 1% of the main diagonal (very similar results are obtained within the range 0.001–5%). Now the resolution remains the same for smaller $\Delta k'$. This can be explained by the fact that with diagonal stabilization, data are not reconstructed in the introduced gap; consequently, the aperture is not larger for smaller $\Delta k'$ and the resolution does not increase.

Instead of diagonal stabilization, other ways of stabilization may be used. Sacchi and Ulrych (1995) use a stabilization method especially focused on improving the resolution.

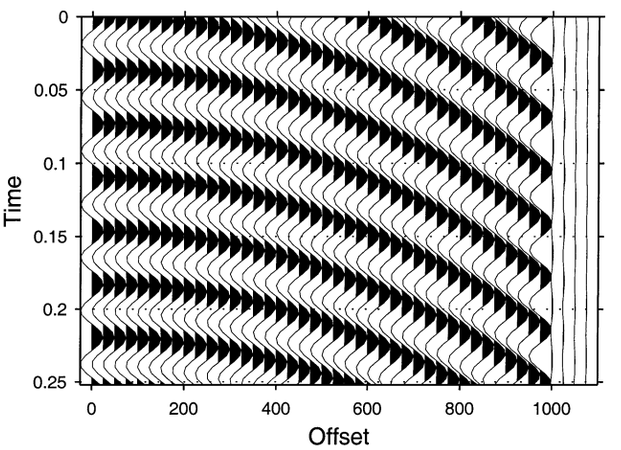

FIG. 11. Reconstructed data with $\Delta k' = 0.75\Delta k'_a$ and diagonal stabilization.

By forcing a certain sparseness in the transform domain, the smearing from the finite aperture can be reduced, giving higher resolution. This makes the method especially suited for filtering multiples that have almost the same curvature as the primaries and for extrapolating offsets. The results of the method depend on how much the events smear in the transform domain from, e.g., AVO effects, and how many events with different curvatures exist in the time window used. On the down side, this method is computationally more expensive; therefore, the parabolic Radon transform with diagonal stabilization will be a more economic choice if the curvature of the multiples is not very close to the curvature of the primaries and when reconstructing missing offsets.

PARABOLIC RADON TRANSFORM WITH FREQUENCY-DEPENDENT ΔQ

If no stabilization is used, then $\Delta q = 2\pi/[\omega(x_{\max}^2 - x_{\min}^2 + \Delta x_a^2)]$ as given in equation (33) is a good choice for the sampling interval in the transform domain.

In combination with diagonal stabilization, the Δq can be made slightly smaller to obtain the best data fit. From a number of numerical experiments we have found that $\Delta q = 2\pi/[\omega(x_{\max}^2 - x_{\min}^2 + 4\Delta x_a^2)]$ is a good choice for the sampling interval (the introduced gap is then four times the actual gap). A further reduction of Δk will not give a better data fit or higher resolution (under the assumption that the maximum q-range is estimated).

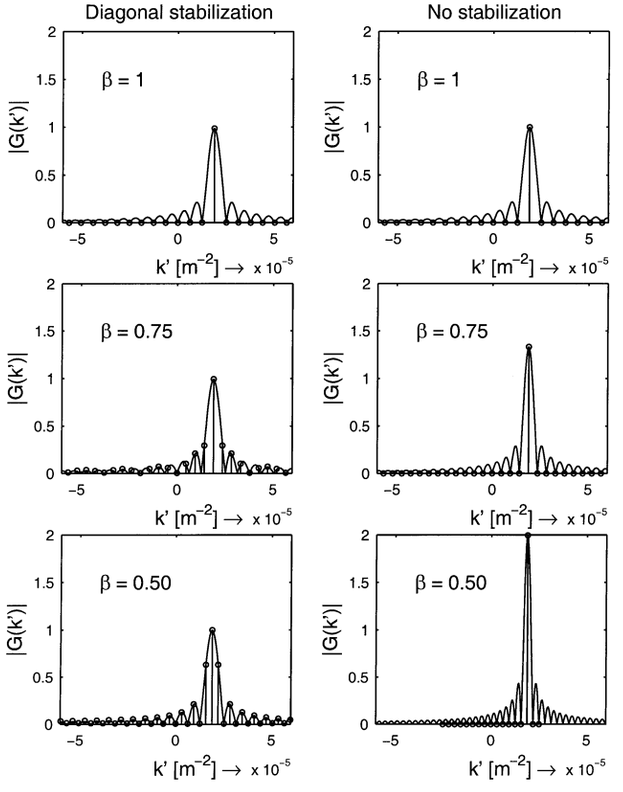

FIG. 12. Response of single curvature event with diagonal stabilization (left) and without stabilization (right) for $\Delta k' = \beta \Delta k'_a = \beta[2\pi/(Y_a + \Delta y_a)]$ [see equation (35)], with $\beta = 1$, 0.75, 0.5 (from top to bottom).

It is interesting to analyze the choice of the sampling interval for the conventional parabolic Radon transform. By using a constant Δq, based on $\omega_{\max}$ [see equation (9)], it follows that for $\omega \ll \omega_{\max}$ the introduced gap becomes very large. If $\omega = 0.5\omega_{\max}$, then this Δq is only approximately half of $\Delta q \propto (1/\omega)$ proposed above, and the introduced gap is larger than the actual aperture. Moreover, the number of q-values to cover the calculated q-range is approximately twice as large as with $\Delta q \propto (1/\omega)$.

Using the parabolic Radon transform with $\Delta q \propto (1/\omega)$ means that (1) for lower frequencies the number of q-values to be calculated is smaller and (2) the complete transform is independent of frequency and the computation of the parabolic Radon transform is equal to the computation of the least-squares nonuniform Fourier transform. As a consequence, computationally more efficient algorithms can be used, as will be shown in the next section. Filtering of multiples with these algorithms is similar to dip filtering in the f-k domain. With these algorithms a transformation to (τ, q) is only possible after an interpolation in the transform domain such that a constant Δq is obtained.

EFFICIENCY

The sequence of forward parabolic Radon transform, filtering in the transform domain, and inverse parabolic Radon transform can be written as

$$\mathbf{d}_f = \mathbf{L}\mathbf{F}(\mathbf{L}^H\mathbf{L} + \alpha^2\mathbf{I})^{-1}\mathbf{L}^H\mathbf{d}, \tag{47}$$

where $\mathbf{F}$ is a diagonal matrix representing the filtering in the transform domain. These computations are done for each temporal frequency; therefore, a forward and inverse Fourier transform along the time direction is required. The total computational costs, however, are dominated by the computations for each temporal frequency.

Now the following major computational steps can be recognized:

1) Direct forward transform $\mathbf{b} = \mathbf{L}^H\mathbf{d}$,
2) Computation of the operator $\mathbf{H} = \mathbf{L}^H\mathbf{L} + \alpha^2\mathbf{I}$,
3) Solving the system $\mathbf{H}\hat{\mathbf{m}} = \mathbf{b}$, and
4) Direct inverse transform $\mathbf{d}_f = \mathbf{L}\hat{\mathbf{m}}$.

The entries of the matrix $\mathbf{H}$ are given by

$$\mathbf{H}_{mn} = \sum_{l=1}^{N} \exp\left(j(m-n)\omega\Delta q x_l^2\right) + \delta_{mn}\alpha^2, \tag{48}$$

where δ_{mn} is the Kronecker delta function. Kostov (1990) recognized the Toeplitz structure of this matrix for the parabolic Radon transform and proposed the Levinson scheme [cost: $O(M^2)$]. For larger systems the Toeplitz-conjugate-gradient scheme (Strang, 1986) can be faster [cost: $O(M \log M)$]. Since the Toeplitz matrix is also Hermitian, one row or column determines the complete matrix. Calculating one row or column is actually a nonuniform discrete Fourier transform.

The direct transforms are also nonuniform discrete Fourier transforms; consequently, the first, second, and fourth steps can efficiently be calculated using the nonuniform fast Fourier

transform (FFT) (see Dutt and Rokhlin, 1993; Beylkin, 1995; and Duijndam and Schonewille, 1999), as proposed by Schonewille and Duijndam (1997) and used by Beylkin (1998), who gives an example where the gain in computation speed is a factor of 3.5.

By using $\Delta q \propto (1/\omega)$, the following advantages are obtained. The Toeplitz operator becomes independent of frequency and is calculated only once. Furthermore, the nonuniform FFT is initialized only once. The computations for each frequency slice are now

1) Direct forward transform: nonuniform FFT, evaluation only,
2) Solving the system $\mathbf{H}\hat{\mathbf{m}} = \mathbf{b}$, and
3) Direct inverse transform: nonuniform FFT, evaluation only.

Since the initialization of the nonuniform FFT costs approximately as much as the evaluation, the gain in efficiency for steps 1 and 3 is approximately a factor of two. Using $\Delta q \propto (1/\omega)$ also improves the efficiency of step 2 since for lower temporal frequencies the number of estimated q-values is smaller (see Figure 3) and therefore the system to be solved is smaller. For step 2, the gain in efficiency is approximately three times for an $O(M^2)$ scheme and more than two times for an $O(M \log M)$ scheme. Because the different steps are at least two times faster, we can conclude that the complete scheme is also at least a factor of two faster than without using $\Delta q \propto (1/\omega)$.

Although with $\Delta q \propto (1/\omega)$, or constant $\Delta k'$, the operator does not directly depend on the frequency, we do vary the k'-range of the reconstruction (see Figure 3). Still, the operator $\mathbf{H}$ must be computed for largest k'-range (highest frequency) only, since for smaller k'-range (lower frequency) simply a part of this matrix can be used.

The k'-range does not change for each temporal frequency slice (except in the case of a very large M); therefore, it is possible to compute the inverse of the matrix once for each k'-range. Inversion of the system (step 2) becomes a multiplication of the inverse matrix with $\mathbf{b}$. This is also an $O(M^2)$ operation but is twice as fast as the Levinson recursion. For larger systems this multiplication can be done more efficiently [cost: $O(M \log M)$] using the Gohberg-Semencul formulations [see Strohmer (1994) and references cited there].

Precomputation of operators

If a number of data gathers with the same geometry must be transformed, then precomputation of operators can be efficient. This method requires precomputation of the forward transform operator $(\mathbf{L}^H\mathbf{L} + \alpha^2\mathbf{I})^{-1}\mathbf{L}^H$ and the inverse transform $\mathbf{L}$. After the operators have been computed, for each frequency slice the only computations are complex matrix vector multiplications [$O(MN)$ operations]. Filtering can be done by applying band-limited forward and backward operators (Kabir and Verschuur, 1993).

Using a $\Delta q \propto (1/\omega)$, the operators do not have to be computed for all frequencies but only for a limited number (using the pie-shaped spectrum, Figure 3). Typically this means that the precomputation of the operator is not the dominant computational step and can be repeated for each data gather, thus giving better quality in case of irregular sampling.

Data reduction

A dramatic efficiency improvement can be obtained if the parabolic Radon transform with $\Delta q \propto (1/\omega)$ is combined with data reduction that can be described as a linear operator. Examples are stacking, selecting a common offset, or filling in missing traces. Equation (47) can be written as $\mathbf{d}_f = \mathbf{T}\mathbf{d}$ with

$$\mathbf{T} = \mathbf{L}\mathbf{F}(\mathbf{L}^H\mathbf{L} + \alpha^2\mathbf{I})^{-1}\mathbf{L}^H. \tag{49}$$

Suppose the data is to be stacked. Since stacking is a summation over the spatial coordinate of $\mathbf{d}_f$, it is equivalent to the inner product with vector $\mathbf{s} = [1, \ldots, 1]^T$:

$$\sum d_f = \mathbf{s}^T\mathbf{d}_f = (\mathbf{s}^T\mathbf{T})\,\mathbf{d}. \tag{50}$$

Now $\mathbf{s}^T\mathbf{T}$ is a row vector, which reduces the combination of parabolic Radon transform and stacking to a weighted stack.

The total cost is a vector inner product for each frequency component plus the calculation of the operator $\mathbf{s}^T\mathbf{T}$, and the Fourier transforms along the time direction. The operator can be calculated very efficiently by noticing that

$$\mathbf{s}^T\mathbf{T} = (\mathbf{T}^H\mathbf{s})^H = (\mathbf{L}(\mathbf{L}^H\mathbf{L} + k^2\mathbf{I})^{-1}\mathbf{F}^H\mathbf{L}^H\mathbf{s})^H, \tag{51}$$

which is similar to the calculations for one frequency in the complete parabolic Radon transform scheme, and can be solved using nonuniform FFTs and any of the Toeplitz solvers. Because $\Delta q \propto (1/\omega)$ is used, the operator is independent of frequency and is calculated only once. The total costs are less than the costs of the complete parabolic Radon transform by an order of magnitude (but of course the complete scheme yields full prestack data).

For selection of a common offset and filling in missing traces, the direct transformation back has only to be done for the relevant offset(s). The total cost is again the calculation of the operator once plus one vector inner product for each reconstructed offset or missing trace for each frequency.

EXAMPLE

The parabolic Radon transform with $\Delta q \propto (1/\omega)$ and the conventional parabolic Radon transform have been applied to an acoustic finite-difference data gather. The laterally invariant subsurface model consists of five layers with velocities [1500, 2000, 2500, 3200, 4200] m/s and density [1000, 2000, 3000, 3300, 3800] kg/m^3. The interfaces are at [300, 700, 1000, 1400] m. In Figure 13 the data are shown after NMO correction and stretch mute. The primaries are at [0.39, 0.79, 1.03, 1.28] s.

The sampling interval in the transform domain is chosen as $\Delta q = 2\pi/[\omega(x^2_{\max} - x^2_{\min} + 4\Delta x^2_a)]$ for the parabolic Radon transform with $\Delta q \propto (1/\omega)$ and $\Delta q = 2\pi/[\omega_{\max}(x^2_{\max} - x^2_{\min} + 4\Delta x^2_a)]$ for the conventional parabolic Radon transform with $\omega_{\max}/2\pi = f_{\max} = 60$ Hz.

For both methods the estimated curvature range corresponds to a moveout at the largest offset from -70 to 160 ms. In principle this leads to a pie shape in the (k', ω) domain (see Figure 3) with a k'-range that is zero at $\omega = 0$. Consequently, only one k'-value would be estimated for the parabolic Radon transform with $\Delta q \propto (1/\omega)$. To improve the handling of low temporal frequencies for the parabolic Radon transform with $\Delta q \propto (1/\omega)$, the curvature range is extended by $2\Delta q$ on both sides (such that for $\omega = 0$, the q-range is $5\Delta q$). It

was verified that the curvature range used was smaller than $\{[(x_{\max}^2 - x_{\min}^2)/\Delta x_a^2] + 2\}\Delta q$.

For the multiple elimination, the primaries between −70 and 20 ms are muted, and the multiples are transformed back to the spatial domain and subtracted from the primaries.

In Figure 14 the data are shown after multiple filtering using the parabolic Radon transform with $\Delta q \propto (1/\omega)$ and in Figure 15 after multiple filtering using the conventional parabolic Radon transform. Because the sampling in the transform domain varies between the methods, the threshold between the primaries and the multiples will also be slightly different for both methods. Usually differences will be very small and not inherent to the method.

The residuals for both methods are calculated by doing a forward and backward transformation without a mute of the primaries and subtracting the result from the original data. In Figure 16 the residuals are shown for the parabolic Radon transform with $\Delta q \propto (1/\omega)$ and in Figure 17 for the conventional parabolic Radon transform. The data are scaled up by a factor of 10. The residuals for the parabolic Radon transform with $\Delta q \propto (1/\omega)$ are clearly smaller, which is mainly because the estimated k'-range using the parabolic Radon transform with $\Delta q \propto (1/\omega)$ is larger than for the conventional parabolic Radon transform. For higher frequencies, the difference between the k'-ranges is relatively small, and the k'-range is large enough for both transforms. For the lowest temporal frequencies, the k'-range for the conventional parabolic Radon transform is very small. It is also possible to enlarge this range for the conventional parabolic Radon transform (except for $\omega = 0$); but because of the very small sampling interval (from the Fourier point of view, see Figure 3), a very large number of parameters must be estimated.

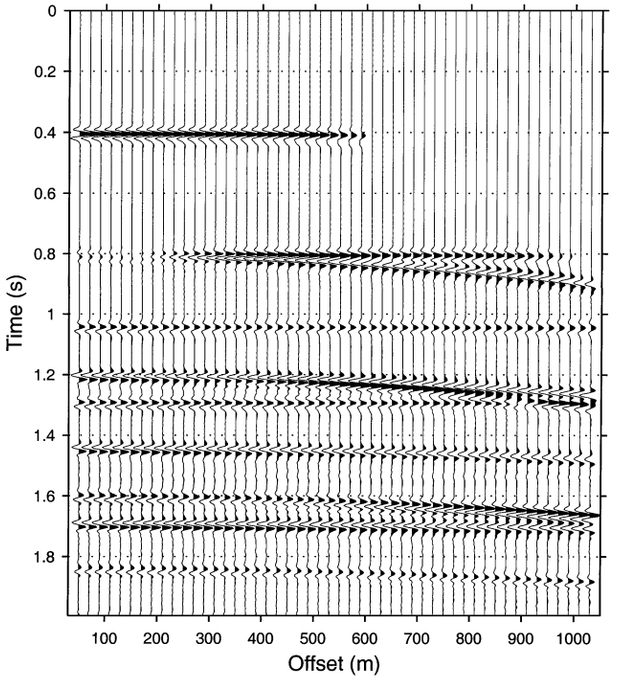

FIG. 13. Finite-difference data after NMO and stretch mute.

CONCLUSION

We have shown that for each temporal frequency component the least-squares parabolic Radon transform is equivalent to the least-squares nonuniform Fourier transform, which provides new and useful insights into the parabolic Radon transform.

Using nonuniform Fourier theory the sampling interval in the transform domain for the parabolic Radon transform has been studied. A critical value for the sampling interval is $\Delta q = 2\pi/[\omega(x_{\max}^2 - x_{\min}^2 + \Delta x_a^2)]$. If Δq is chosen larger, then wraparound and edge effects occur and the data fit is decreased. If Δq is chosen significantly smaller, then stabilization is required. In combination with diagonal stabilization, Δq can be reduced slightly for the best data fit, and we propose to use $\Delta q = 2\pi/[\omega(x_{\max}^2 - x_{\min}^2 + 4\Delta x_a^2)]$. A smaller Δq does not yield a higher resolution or better data fit and leads to a larger matrix to be inverted for the same curvature range.

The theoretical maximum curvature range that can be estimated is $N\Delta q$, but stability will decrease beyond $\{[(x_{\max}^2 - x_{\min}^2)/\Delta x_a^2] + 2\}\Delta q$. Estimating a significantly larger q-range while using diagonal stabilization to handle instability leads to smearing of energy. This smearing of energy is caused when the reconstructed signal is forced to zero between the traces at large offsets. If the offset coordinate is sparsely sampled, then by using a limited offset range for higher temporal frequencies the q-range can be enlarged without instability problems.

The efficiency of the parabolic Radon transform can be improved by using a $\Delta q \propto (1/\omega)$, which makes the complete transform independent of frequency, and by using the nonuniform FFT. Using $\Delta q \propto (1/\omega)$, operators can be precomputed for each gather, taking into account the exact locations of the samples and therefore yielding better quality with only a very small increase of computational costs. If the parabolic Radon transform with $\Delta q \propto (1/\omega)$ is combined with data reduction, very fast schemes are possible.

Using $\Delta q \propto (1/\omega)$ means that the data cannot be transformed directly to (τ, q); an interpolation in the Radon domain is required.

The parabolic Radon transform with $\Delta q \propto (1/\omega)$ has been compared with the conventional parabolic Radon transform in a synthetic example. As expected, the quality of the multiple elimination is approximately equal, where the residuals of the parabolic Radon transform with $\Delta q \propto (1/\omega)$ are smaller thanks to a larger curvature range for low frequencies.

ACKNOWLEDGMENTS

We thank Prof. J. T. Fokkema, Dr. D. J. Verschuur, Dr. J. Brittan, Dr. P. W. Cary, and the anonymous reviewers for their valuable comments and helpful discussions.

REFERENCES

Bagchi, S., and Mitra, S. K., 1996, The nonuniform discrete Fourier transform and its applications in filter design: Part I—1-D: IEEE Trans. Circuits Syst. II, **43**, 422–433.

Beylkin, G., 1995, On the Fast Fourier transform of functions with singularities: Appl. Comp. Harm. Anal., **2**, 363–381.

——— 1998, Fast Radon transform for multiple elimination: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1351–1352.

Cary, P. W., 1998, The simplest discrete Radon transform: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1999–2002.

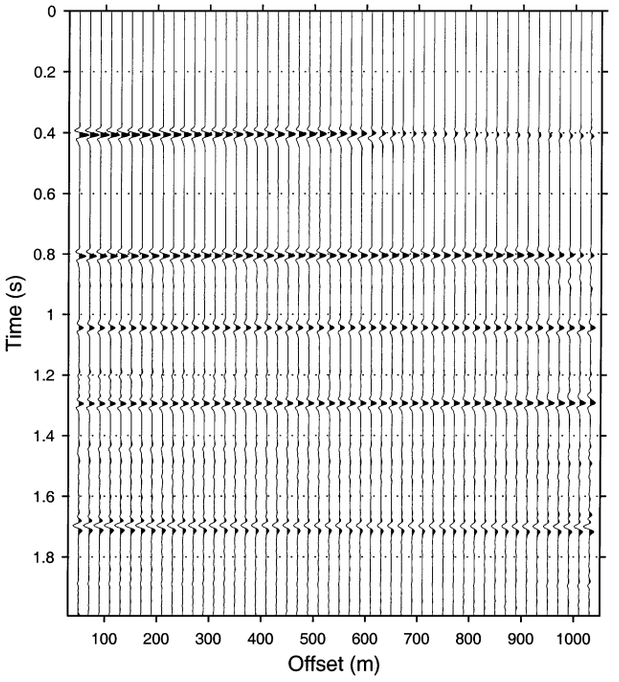

FIG. 14. Data after multiple filtering with the parabolic Radon transform with $\Delta q \propto (1/\omega)$.

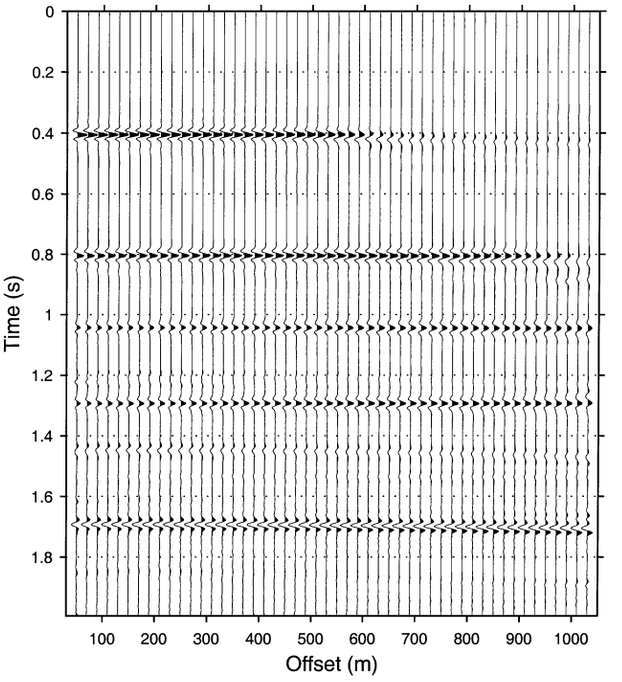

FIG. 15. Data after multiple filtering with the conventional parabolic Radon transform.

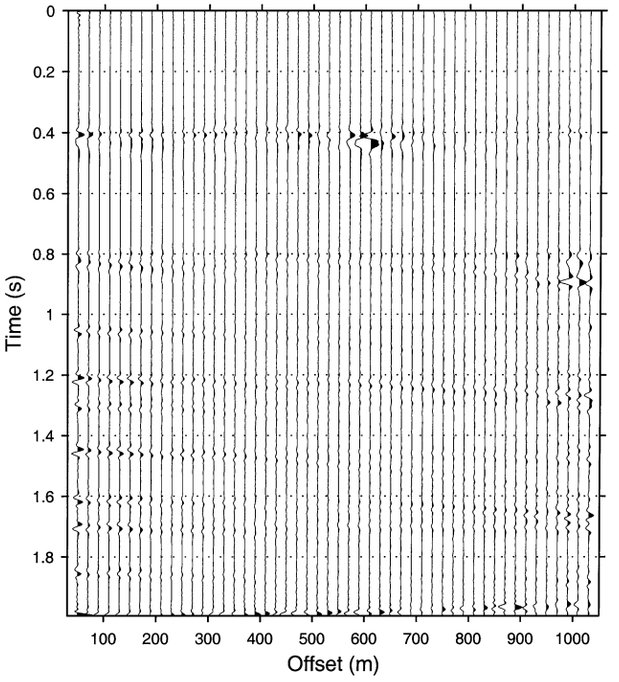

FIG. 16. Residuals for the parabolic Radon transform with $\Delta q \propto (1/\omega)$, scaled up by a factor of 10.

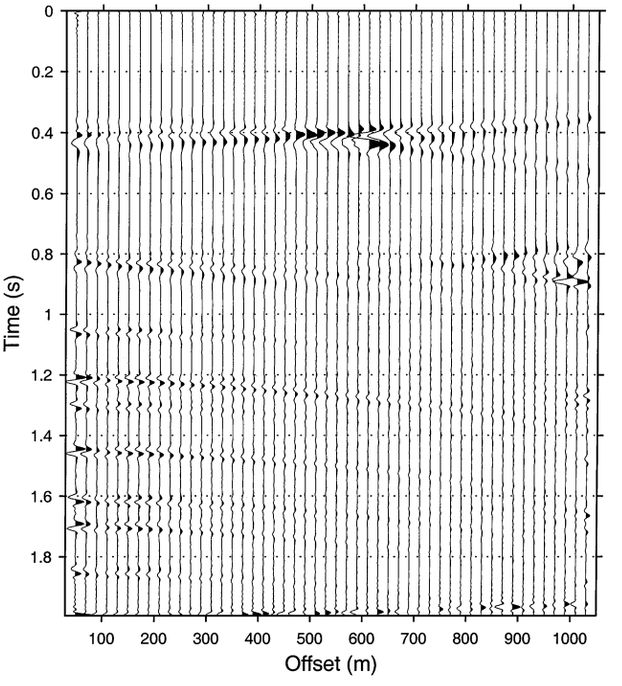

FIG. 17. Residuals for the conventional parabolic Radon transform, scaled up by a factor of 10.

Duijndam, A. J. W., and Schonewille, M. A., 1999, Nonuniform Fast Fourier transform: Geophysics, **64**, 539–551.
Duijndam, A. J. W., Schonewille, M. A., and Hindriks, C. O. H., 1999, Reconstruction of band-limited signals, irregularly sampled along one spatial direction: Geophysics, **64**, 524–538.
Dunne, J., and Beresford, G., 1995, A review of the τ–p transform, its implementation and its applications in seismic processing: Expl. Geophys., **26**, 19–36.
Dutt, A., and Rokhlin, V., 1993, Fast Fourier transforms for nonequispaced data: SIAM J. Sci. Comp., **14**, 1368–1393.
Feichtinger, H. G., Gröchenig, K., and Strohmer, T., 1995, Efficient numerical methods in non-uniform sampling theory: Numerische Mathematik, **69**, 423–440.
Gröchenig, K., 1993, A discrete theory of irregular sampling: Linear Algebra and Its Applications: **193**, 129–150.
Gulunay, N., 1990, *F-x* domain least-squares tau-p and tau-q: 60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1607–1610.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: J. Can. Soc. Expl. Geophys., **22**, 44–55.
Hugonnet, P., and Canadas, G., 1995, Aliasing in the parabolic Radon transform: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1366–1369.
Kabir, M. M. N., and Verschuur, D. J., 1993, Parallel computing of the parabolic Radon transform—Application for CMP-based preprocessing: 63rd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 193–196.
Kelamis, P. G., and Chiburis, E. F., 1992, Land data examples of Radon multiple suppression: First Break, **10**, 275–280.
Kostov, C., 1990, Toeplitz structure in slant-stack inversion: 60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1618–1621.
Marfurt, K. J., Schneider, R. V., and Mueller, M. C., 1996, Pitfalls of using conventional and discrete Radon transforms on poorly sampled data: Geophysics, **61**, 1467–1482.
Sacchi, M. D., and Ulrych, T. J., 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169–1177.
Schonewille, M. A., and Duijndam, A. J. W., 1997, Parabolic Radon transform and x-squared fk-transform—Aliasing and efficiency: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1410–1413.
Strang, G., 1986, A proposal for Toeplitz matrix calculations: Studies in Appl. Math., **74**, 171–176.
Strohmer, T., 1994, Recovery of missing segments and lines in images: Optic. Eng. SPIE 1994, 3283–3289.
Zhou, B., and Greenhalgh, S. A., 1994, Linear and parabolic τ-p transforms revisited: Geophysics, **59**, 1133–1149.