*Geophysical Prospecting*, 1998, **46**, 303–322

# Filtering vibroseis data in the precorrelation domain[1]

Shuang Qin[2,3] and David K. Smythe[2]


## Abstract

Vibroseis data recorded at short source–receiver offsets can be swamped by direct waves from the source. The signal-to-noise ratio, where primary reflections are the signal and correlation side lobes are the noise, decreases with time and late reflection events are overwhelmed. This leads to low seismic resolution on the vibroseis correlogram. A new precorrelation filtering approach is proposed to suppress correlation noise. It is the 'squeeze-filter-unsqueeze' (SFU) process, a combination of 'squeeze' and 'unsqueeze' (S and U) transformations, together with the application of either an optimum least-squares filter or a linear recursive notch filter. SFU processing provides excellent direct wave removal if the onset time of the direct wave is known precisely, but when the correlation recognition method used to search for the first arrival fails, the SFU filtering will also fail. If the tapers of the source sweeps are badly distorted, a harmonic distortion will be introduced into the SFU-filtered trace. SFU appears to be more suitable for low-noise vibroseis data, and more effective when we know the sweep tapers exactly. SFU requires uncorrelated data, and is thus cpu intensive, but since it is automatic, it is not labour intensive.

With non-linear sweeps, there are two approaches to the S,U transformations in SFU. The first requires the non-linear analytical sweep formula, and the second is to search and pick the zero nodes on the recorded pilot trace and then carry out the S,U transformations directly without requiring the algorithm or formula by which the sweep was generated. The latter method is also valid for vibroseis data with a linear sweep. SFU may be applied to the removal of any undesired signal, as long as the exact onset time of the unwanted signal in the precorrelation domain is known or determinable.


## Introduction

Vibroseis reflection data recorded at short source–receiver offsets can be swamped by direct waves from the source. After correlation, we usually think of the central peak and

---

[1] Paper presented at the 57th EAGE Conference – Geophysical Division, Glasgow, UK, May–June 1995. Received April 1997, revision accepted November 1997.
[2] Department of Geology & Applied Geology, University of Glasgow, Glasgow G12 8QQ, UK.
[3] Present address: Ulterra Geoscience Ltd, 715 5th Avenue SW, Suite 520, Calgary, Alberta T2P 2X6, Canada.

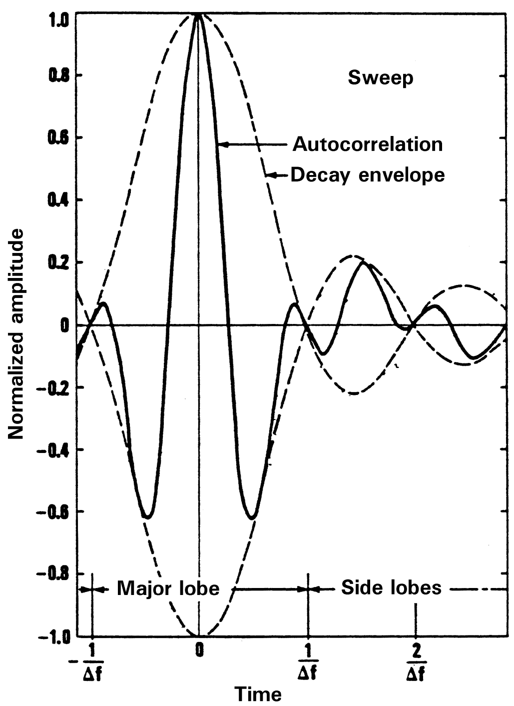



**Figure 1.** Autocorrelation of a sweep (after Cunningham 1979), showing the definition of major lobe and side lobes.

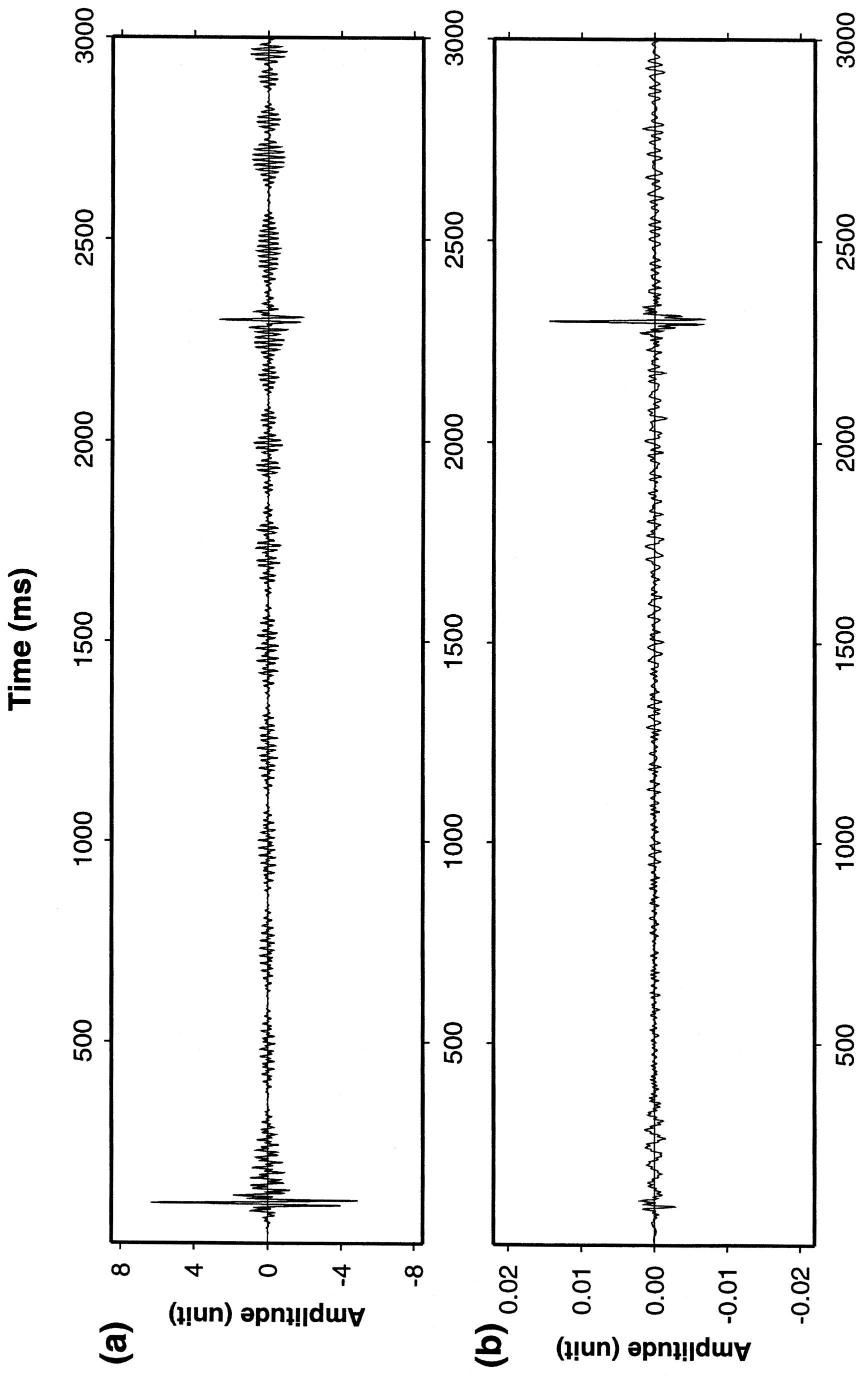
adjacent half-cycles as being so large in comparison to the amplitude of the later side lobes that the side lobes can be ignored. Side lobes can be considered to be undesired signal-generated noise since their form is signal dependent. Naturally, we want the side lobes to be as small as possible so as not to obscure nearly coincident and/or weaker signals. The first side lobe of a conventional linear sweep can be as high as 22% of the peak and the second side lobe can be as high as 13% of the peak (Fig. 1). These side lobes will interfere with the correlation peaks of desired events arriving at the same times as the unwanted side lobes. The task of resolving the interfering events is usually carried out by spiking deconvolution, phase shifting and/or spectral whitening techniques (e.g. Yilmaz 1987), which are not the subject of this paper.

However, at time ranges outside the first couple of side lobes shown in Fig. 1, the decay envelope of the side lobes flattens out, so that the far outlying side lobes from first arrivals, for example, can still be as high in amplitude as the major lobes of desired reflection events. The result is that the signal-to-noise ratio, where the correlation peaks of primary reflections are the signal and correlation side lobes are the noise, decreases with time, and later events can be overwhelmed. The problem is one of correlation noise. The synthetic example in Fig. 2(a) shows that side lobes from a large event such as the first break can dominate much later, low-amplitude signals. The sweep is a 10–60 Hz linear sweep with 0.5 s linear tapers. The amplitude of the first signal at 100 ms (the first break) is 4000 times larger than that of the second signal (reflection) arriving 2200 ms later. An exponential gain has been applied to the correlated trace to make the two signals visible; this gain makes the first break look asymmetrical. The reflection signal correlation peak is 2.7 times bigger than the side-lobe contamination in its vicinity. We define this ratio as the signal-to-noise (S/N) ratio.

Correlation-noise suppression has been discussed since the introduction of correlation. Klauder *et al.* (1960) described measures to reduce side-lobe levels by frequency weighting, and the sweep-tapering technique has become routine for correlation-noise suppression since the early 1960s. A considerable reduction of the side-lobe noise can be achieved by tapering the amplitude spectrum of the vibrator sweep, at the expense of loss of energy put into the ground (Edelmann 1966). During the 1970s, the Combisweep technique was introduced by Werner and Krey (1979), and was discussed as a correlation-noise suppression technique by Edelmann and Werner (1982). This technique starts from the idea that neither a non-linear sweep nor amplitude shaping can cope with the coupling and attenuating effects of the ground, which influence the vibroseis signal. The Combisweep technique yields an improved recorded signal by strengthening certain spectral components which otherwise would remain below the noise level.

At about the same time, coding techniques to suppress correlation noise were introduced from communication theory into vibroseis exploration. One of them is the pseudorandom coding technique developed by Cunningham (1979). This encoding, however, makes great demands upon the transient response of the vibrator. Another coding technique was described by Bernhardt and Peacock (1978), designed to suppress far-correlation noise. However, near-correlation noise is only moderately reduced.

All the approaches mentioned above attack the correlation-noise problem either at source (by designing better sweeps), or during post-correlation processing. Our approach is to remove the events causing the problem before correlation. These events will usually be large-amplitude events such as first breaks, ground roll and air blast. The synthetic example shown in Fig. 2(a) has had the first break removed before correlation (Fig. 2b). This output trace is displayed at a constant gain. The amplitude of the desired reflection at 2300 ms has been increased by about 18 000 relative to the first break at 100 ms, and there is a 3.2 increase in the S/N ratio in the vicinity of the reflection event. The first break was removed with our 'squeeze-filter-unsqueeze' algorithm (SFU for short), which is a combination of squeeze and unsqueeze transformations, together with the application of either an optimum least-squares (OLS) filter or a linear recursive notch filter prior to the inverse transform. We squeeze the uncorrelated trace in a non-linear way so that the unwanted signal is transformed into a constant quasi-sinusoidal wave of known phase and duration. We then filter this signal out, without removing any other signal components. Lastly, we unsqueeze the trace. SFU is a single-trace process applied to uncorrelated data. The following sections present the description of the new filter, its implementation, and a discussion


Time (ms)
(a)
Amplitude (unit)
(b)
Amplitude (unit)

of the algorithm performance examined through use of both synthetic modelling data and field data examples.

## The problem of side lobes

In order to demonstrate an example of the correlation side-lobe problem, we examine the conventional linear swept-frequency sine wave from start frequency $f_1$ to end frequency $f_2$ over the duration $T$. Its autocorrelation function $\Phi(\tau)$ can be well approximated for seismic applications (Cunningham 1979) by

$$\Phi(\tau) = \frac{A^2}{2} \frac{T \sin \pi \Delta \mathrm{f} \tau}{\pi \Delta f \tau} \cos 2\pi \left( \mathrm{f}_0 + \frac{\Delta \mathrm{f} \tau}{2\mathrm{T}} \right) \tau, \text{ for } 0 \leq \tau \leq \mathrm{T}, \tag{1}$$

where $\tau$ denotes processed record time, $A$ denotes signal amplitude, $\Delta f = f_2 - f_1$ is the bandwidth, and $f_0 = (f_2 + f_1)/2$ is the centre frequency.

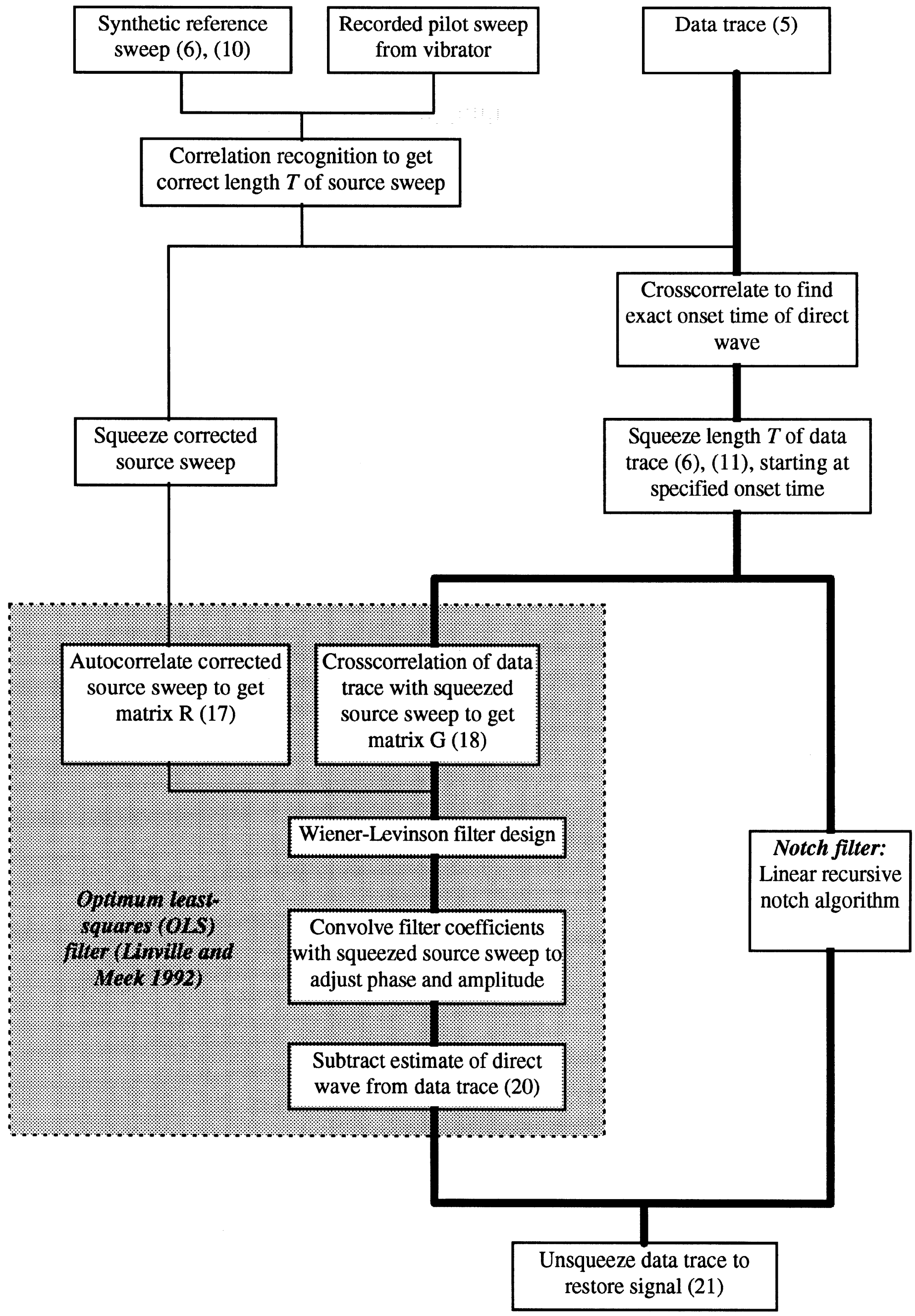
The autocorrelation is plotted in Fig. 1 as the solid curve, which represents the product of a constant, a sin $x/x$ decay envelope (where $x = \pi \Delta f \tau$; the dashed curve), and a cosine term. The function $\Phi(\tau)$ has zero amplitude values at integer multiples of $1/\Delta f$, the reciprocal of the sweep bandwidth, and oscillates within the decay envelope at a rate governed by the arguments of the cosine term. The portion of the autocorrelation between $-1/\Delta f$ and $1/\Delta f$ is defined as the major lobe and that portion outside these points as the side lobes. The side lobes are undesired signal-generated noise since their form is signal dependent, and, as we see from the synthetic example in Fig. 2(a), they resemble coherent noise which can dominate later arrived signals. Naturally, we want the side lobes to be as small as possible so as not to obscure nearby and/or weaker signals.

The principle of SFU filtering is based on the fact that the side lobes caused by strong direct waves may be the dominant coherent noise problem. If the direct wave can be cancelled in the precorrelation domain, this source of correlation noise can be removed without attenuating the signal, so that the signal-to-noise ratio will be improved as shown in Fig. 2(b).

## The SFU filter algorithm

Figure 3 shows the flow chart of the SFU algorithm. The squeezing and unsqueezing transformations are the fundamental components of SFU. Figure 4 illustrates their operation on a synthetic linear sweep. Firstly, a forward transformation linearly squeezes the portion of the trace from the beginning of the undesired signal (e.g. the direct wave) to the end of it (Fig. 4a), to turn the direct wave into a constant-frequency

**Figure 2.** Correlated synthetic data trace from a 10–60 Hz linear sweep with 500 ms linear tapers. The direct wave is at 100 ms and the reflection is at 2300 ms, with 1/4000 of the amplitude of the direct wave. (a) Before SFU; spherical divergence gain applied; (b) after removal of the direct wave by SFU, constant gain applied.


Synthetic reference sweep (6), (10)
Recorded pilot sweep from vibrator
Data trace (5)
Correlation recognition to get correct length T of source sweep
Crosscorrelate to find exact onset time of direct wave
Squeeze corrected source sweep
Squeeze length T of data trace (6), (11), starting at specified onset time
Autocorrelate corrected source sweep to get matrix R (17)
Crosscorrelation of data trace with squeezed source sweep to get matrix G (18)
Wiener-Levinson filter design
Optimum least-squares (OLS) filter (Linville and Meek 1992)
Convolve filter coefficients with squeezed source sweep to adjust phase and amplitude
Subtract estimate of direct wave from data trace (20)
Notch filter: Linear recursive notch algorithm
Unsqueeze data trace to restore signal (21)

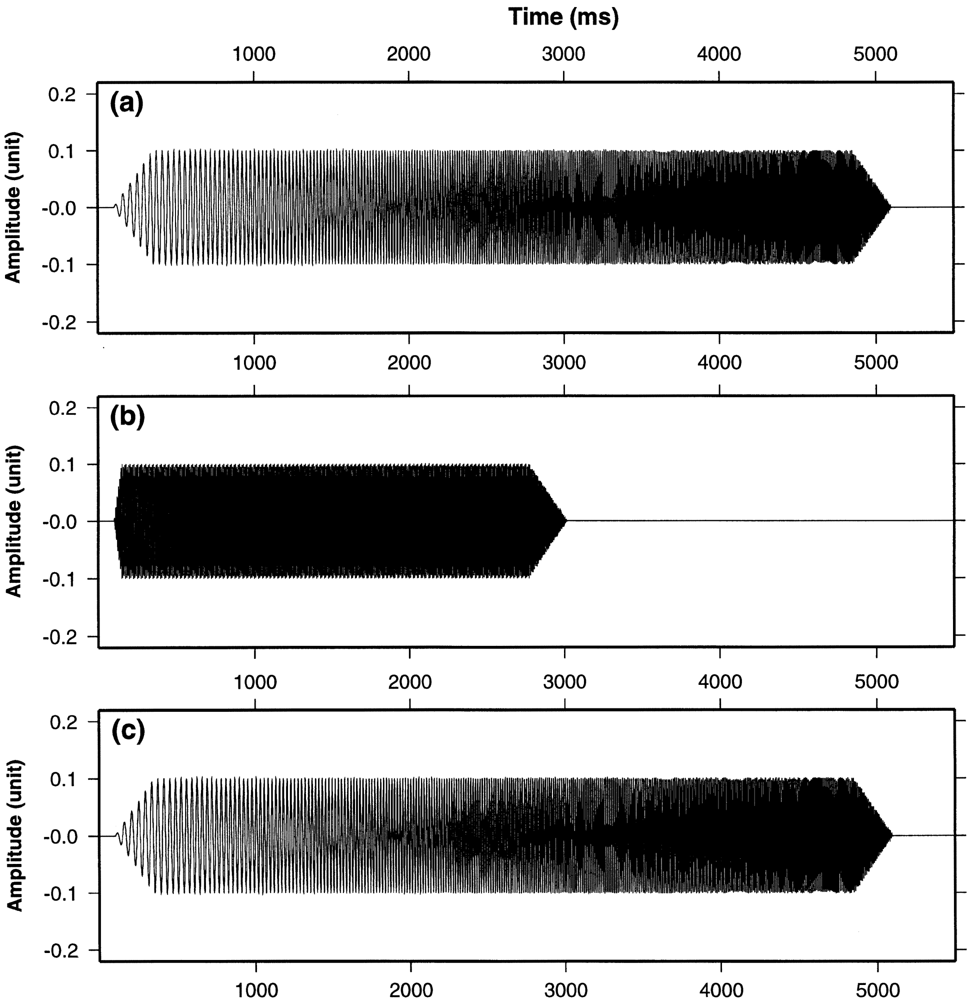



**Figure 4.** Demonstration of the squeezing and unsqueezing of a sweep: (a) a linear sweep from 10 Hz to 60 Hz over 5 s; (b) squeezed to a constant 60 Hz; (c) the unsqueezed (recovered) sweep from (b).

signal (Fig. 4b). A suitable filter can then remove it. The resulting output of such a filter is not shown; instead we show that the inverse transformation (unsqueezing) successfully restores the original trace (Fig. 4c). Since vibroseis data are recorded with sources using both linear and non-linear sweeps, different approaches to the

---

**Figure 3.** Flow chart of the squeeze-filter-unsqueeze (SFU) filter algorithm. Equations in the text are referred to by the numbers within brackets. Heavy linking lines show trace data flow paths; light lines show reference sweep flow paths. The shaded area shows the algorithm of Linville and Meek (1992).

squeezing and unsqueezing transformations have been developed separately for each, and are discussed below.

*Squeezing a linear sweep*

Generally the vibroseis direct wave can be considered as a copy of the source sweep, which in this case takes the form of a sinusoidal linear upsweep ranging from start frequency $f_1$ to end frequency $f_2$, with a duration $T$, at a constant amplitude $A$, and is given by

$$s(t) = A \sin 2p(f_1 + Kt)t, \tag{2}$$

where

$$K = \frac{f_2 - f_1}{2T}. \tag{3}$$

Substituting (3) into (2), we get

$$s(t) = A \sin \frac{\pi}{T}\left[2Tf_1 t + (f_2 - f_1)t^2\right]. \tag{4}$$

We assume that the data trace $x(t)$ comprises the vibroseis direct wave $s(t)$ and signal $y(t)$, e.g. the reflections, so that

$$x(t) = s(t) + y(t). \tag{5}$$

The data trace after the squeezing transform should be

$$\begin{aligned} x^1(t_1) &= S\{s(t) + y(t)\} = s^1(t_1) + y^1(t_1), \\ &= A \sin 2\pi f_0 t_1 + y^1(t_1), \end{aligned} \tag{6}$$

where $S$ denotes the squeezing transform, $s^1(t_1)$ is the squeezed direct-wave sweep with constant frequency $f_0$, $y^1(t_1)$ is the squeezed version of $y(t)$ and

$$t_1 = \frac{\left[2Tf_1 + (f_2 - f_1)t\right]t}{2Tf_0}. \tag{7}$$

*Squeezing a non-linear sweep*

The vibroseis source sweep (and therefore the direct wave) may alternatively be defined as a sinusoidal non-linear upsweep ranging from start frequency $f_1$ to end frequency $f_2$, with a duration $T$, at a constant amplitude $A$, where, for example,

$$s(t) = A \sin 2p(f_1 + K\sqrt{t})t, \tag{8}$$

$$K = \frac{2(f_2 - f_1)}{3\sqrt{T}}. \tag{9}$$

Substituting (9) into (8), we get

$$s(t) = A \sin \frac{2\pi}{T}\left[Tf_1 t + 2(f_2 - f_1)\sqrt{Tt^3}\right]. \tag{10}$$

Similarly, we obtain the data trace after the squeezing transform:

$$x^1(t_1) = S\{s(t) + y(t)\} = s^1(t_1) + y^1(t_1),$$
$$= A \ \sin 2\pi f_0 t_1 + y^1(t_1), \tag{11}$$

where $S$ denotes the squeezing transform, $s^1(t_1)$ is the squeezed direct-wave sweep with constant frequency $f_0$, $y^1(t_1)$ is the squeezed version of $y(t)$, and

$$t_1 = \frac{[Tf_1 + 2(f_2 - f_1)\sqrt{Tt}]t}{Tf_0}. \tag{12}$$

*Filtering methods*

In the following discussion, the squeezing transform is carried out by setting $f_0 = f_2$. In other words the uncorrelated direct-wave signal is squeezed, to turn it into a constant sinusoidal signal with a frequency equal to the upper limit of the vibroseis source sweep; however, this particular choice of frequency is not an essential feature of the transformation.

In order to remove the squeezed constant-frequency sweep $s^1(t_1)$, either a notch filter or an optimum least-squares (OLS) filter can be applied. In the application of a notch filter, a linear recursive notch algorithm is used to remove the narrow frequency band around the squeezed fiducial frequency $f = f_2$ . The output trace sequence $y_n^1$ of the notch filter for an input trace sequence $x_n^1$ (Press *et al.* 1989) is

$$y_n^1 = \sum_{k=0}^{M} C_k x_{n-k} + \sum_{j=1}^{N} b_k y_{n-j}. \tag{13}$$

In (13), the $M + 1 = 3$ filter coefficients $c_k$ and the $N = 2$ coefficients $b_k$ are used here in the form below, and are obtained from the response function:

$$c_0 = \frac{1 + f_2^2}{(1 + \theta f_2)^2 + f_2^2},$$

$$c_1 = -2\frac{1 - f_2^2}{(1 + \theta f_2)^2 + f_2^2},$$

$$c_2 = \frac{1 + f_2^2}{(1 + \theta f_2)^2 + f_2^2},$$

$$b_1 = 2\frac{1 - \theta f_2^2 - f_2^2}{(1 + \theta f_2)\}^+ f_2^2},$$

$$b_2 = \frac{(1 - \theta f_2)^2 + f_2^2}{(1 + \theta f_2)^2 + f_2^2},$$

where the parameter $\theta$ is the desired width of the notch.

In the application of an OLS filter (Ensing 1983; Linville and Meek 1992) to remove the squeezed constant frequency signal from (6) and (11), the squeezed data trace is assumed to comprise the useful signals caused by reflectivity and the squeezed vibroseis direct-wave signal $s_t^1$, which is defined as a sinusoidal noise at the frequency $f_2$ with a constant amplitude $A$ and tapers on both ends.

The aim of the OLS filter is to find the optimum Wiener–Levinson filter coefficients, which are then used to adjust the amplitude and phase of a squeezed reference sweep to match those of squeezed direct-wave sweep $s_t^1$. When the onset time of the first arrival is accurately known, the squeezed synthetic reference sweep is defined as

$$h_t = \sin 2\pi \cdot f_2 t,$$

where $f_2$ is the terminal frequency of the vibroseis direct-wave sweep.

In order to match the amplitude of the squeezed direct-wave signal $s_t^1$ in the data trace $x_t^1$, $h_t$ must be adjusted by convolution with a filter $\nu_t$, so that

$$h_t^1 = \nu_t * h_t. \tag{14}$$

Usually the direct-wave amplitude in the data trace is larger than any other signal, and the least-squares error $L$ between the direct-wave and reference sweep is defined as (Wiener 1949; Yilmaz 1987)

$$L = \sum_t (s_t^1 - h_t^1)^2. \tag{15}$$

Substituting (14) into (15), we obtain

$$L = \sum_t \left( s_t^1 - \sum_\tau \nu_t h_{t-\tau} \right)^2.$$

We wish to compute the filter coefficients ($\nu_0$,$\nu_1$, $\nu_2$, ...) so that the error $L$ is a minimum. Taking the partial derivatives of $L$ with respect to the filter coefficients $\nu_j$ and setting them equal to zero, we get

$$\frac{\partial L}{\partial \nu_j} = -2\sum_t s_1^1 h_{t-i} + 2\sum_t \left( \sum_\tau \nu_t h_{t-\tau} \right) h_{t-j} = 0$$

or

$$\sum_t \nu_t \sum_\tau h_{t-\tau} h_{t-j} = \sum_t s_t^1 h_{t-j},\ j = 0, 1, 2, \ldots, (n-1). \tag{16}$$

By using

$$\sum_t h_{t-\tau} h_{t-j} = r_{j-\tau} = R \tag{17}$$

and

$$\sum_t s_t^1 h_{t-j} = g_j = G, \tag{18}$$

equation (15) becomes

$$\begin{bmatrix} r_0 & r_1 & r_2 & \dots r_{n-1} \\ r_1 & r_2 & r_3 & \dots r_{n-2} \\ \cdot & \cdot & \cdot & \dots \cdot \\ r_{n-1} & r_{n-2} & r_{n-3} & \dots r_0 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \cdot \\ v_{n-1} \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \\ \cdot \\ g_{n-1} \end{bmatrix}. \tag{19}$$

This Toeplitz matrix can be solved rapidly using Levinson's recursion algorithm.

The data trace after the suppression of the direct wave is obtained by subtracting the corrected, squeezed and filtered synthetic sweep:

$$y_t^1 = s_t^1 - v_t * h_t. \tag{20}$$

Due to the cyclic nature of the squeezed sweeps on both the recorded reference and the data trace, a short filter length generally satisfies the requirement of the processing.

*Unsqueezing the filtered trace*

After the direct wave has been removed, the inverse transformation is applied to unsqueeze linearly the squeezed portion $y^1(t)$ of the recorded data trace. This restores the signal deformed by the squeeze transformation. According to (5) and (6), the recovered data trace $y(t)$ becomes

$$y(t) = S^{-1}\{y^1(t)\} = y^1(t_1), \tag{21}$$

where $S^{-1}$ denotes the unsqueezing transformation. For a linear sweep,

$$t_1 = \frac{\sqrt{T^2 f_1^2 + (f_2 - f_1) f_2 t} - T f_1}{f_2 f_1},$$

and for a non-linear sweep, $t_1$ can be obtained by the corresponding squeeze transformation, since $t_1$ cannot be represented as an explicit function of $t$.

Figure 4 demonstrates the operation of the squeezing and unsqueezing transformations without the intervening filter operation, to show that there is no distortion for a synthetic linear upsweep when the onset time of the sweep is known. When the pilot sweep cannot be described analytically, or it is otherwise impossible for it to be described as a function of time, we can still carry out the squeezing and unsqueezing transformations directly by searching and picking the zero nodes on the recorded pilot sweep trace, without having to derive a transformation algorithm. This method can also be applied to vibroseis data recorded with a sweep describable analytically.

*Implementation of SFU*

Because the squeezing and unsqueezing transformations of SFU must be performed from the start of the uncorrelated first break, the exact onset time of this unwanted signal must be determined. Due to static delays, the onset time of the direct wave usually varies from trace to trace in a somewhat irregular way, i.e. it cannot be specified

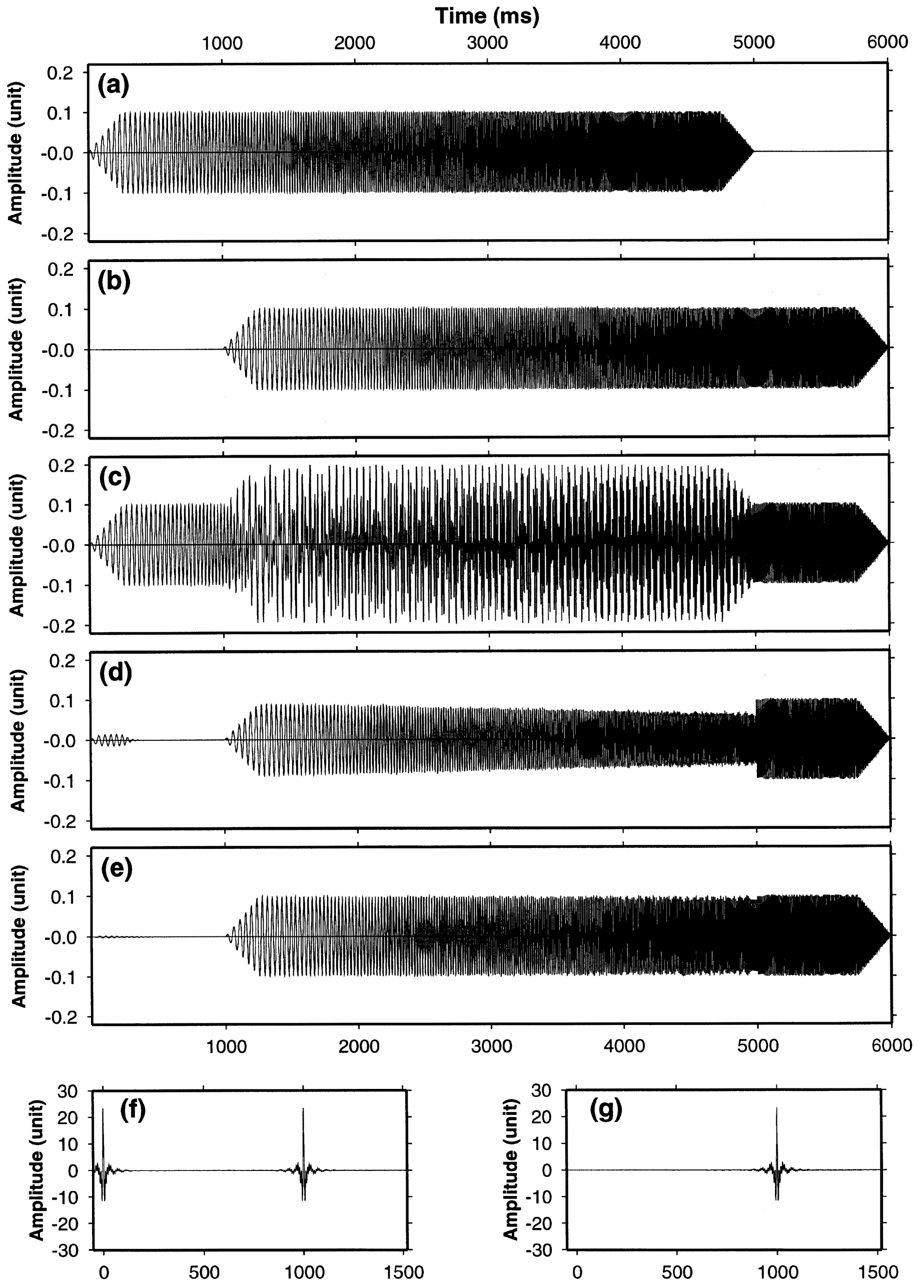



**Figure 5.** Synthetic trace made up of (a) direct arrival at time 0 ms, amplitude 0.1 unit, and (b) reflection arriving at 1000 ms, amplitude 0.1, giving (c) the recorded trace. The direct wave is removed by application of either (d) an optimum least-squares filtering algorithm, or (e) a linear recursive notch algorithm. The correlated trace is shown (f) before application of SFU and (g) after application of SFU.

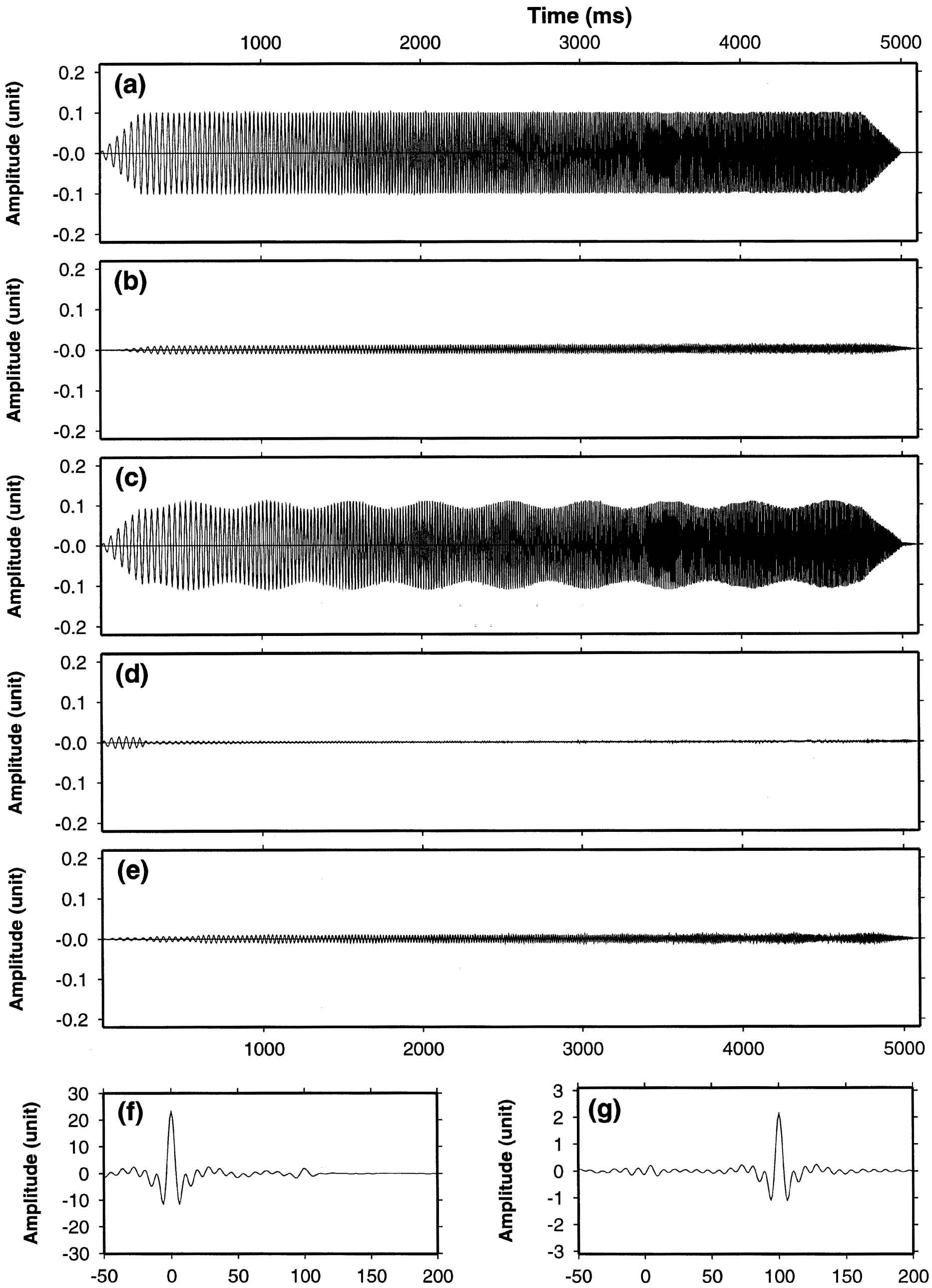



**Figure 6.** Synthetic trace made up of (a) direct arrival at time 0 ms, amplitude 0.1 unit, and (b) reflection arriving at 100 ms, amplitude 0.1, giving (c) the recorded trace. The direct wave is removed by application of either (d) an optimum least-squares filtering algorithm, or (e) a linear recursive notch algorithm. The correlated trace is shown (f) before application of SFU and (g) after application of SFU.

simply as a function of head-wave velocity and offset. To circumvent this problem, an automatic method based on cross-correlation analysis has been developed to find the actual start time. A synthetic sweep defines the onset and end time of the recorded pilot sweep from the vibrator to correct the sweep length *T*. We then correlate the corrected reference source sweep with each data trace individually to determine the onset time of the direct wave for that data trace.

## Synthetic examples

In this section, two examples are presented to illustrate and evaluate SFU. The synthetic upsweep representing the direct wave ranges from 10 to 60 Hz, 5 s duration, with amplitude of 0.1 (arbitrary units) and with 0.5 s linear tapers on both ends (Figs 5a, 6a). All records are displayed with the same gain for comparison. Six seconds of data were generated at a sample interval of 2 ms.

In the first example, an uncorrelated reflection event is generated using the same upsweep as the direct wave, with the same amplitude of 0.1, but delayed by 1 s, as shown in Fig. 5(b). The synthetic seismogram is the superimposition of the direct wave and this reflection signal (Fig. 5c). We then apply SFU to this synthetic trace in two different ways, producing the filtered records shown in Fig. 5(d,e). Figure 5(d) is the output of SFU with the application of an optimum least-squares (OLS) filter, and Fig. 5(e) is the output of the SFU with the application of a linear recursive notch filter. The correlated versions both before and after the application of SFU are shown in Fig. 5(f,g), respectively. SFU has effectively removed the uncorrelated direct wave from the synthetic trace. However, this example is a somewhat limited test, in that the two events (Fig. 5a,b) are well separated in time and are of equal amplitude.

We show next a more stringent test in which the two signals are very close together, and in which the desired signal is also much smaller than the undesired signal. Figure 6 shows the same uncorrelated direct wave (Fig. 6a), but with a 10 times smaller amplitude for the reflection (Fig. 6b) than in Fig. 5. The reflection event is only 100 ms behind the undesired direct wave. These two events are summed to give the synthetic trace (Fig. 6c). Application of SFU successfully removes the uncorrelated direct wave, whichever of the two methods of applying it is used. Figure 6(d) shows the output of SFU with the OLS filter, and Fig. 6(e) shows the output of SFU using the notch filter. The significant improvement in the signal-to-noise ratio is illustrated by the comparison of the correlated versions before and after the removal of the direct wave (Fig. 6f,g, respectively). The strong direct wave has been removed in the precorrelation domain without attenuating the nearby signal. The example in Fig. 2 also shows that the weaker signals obscured by far-correlation noise can be much improved after filtering in the precorrelation domain.

If the tapers of the source sweep are known precisely, SFU with the OLS filter causes less distortion of the uncorrelated reflection than with the notch filter, but when the tapers of the reference sweep do not accurately match those of the direct wave, for whatever reason, we find that harmonic noise results. To illustrate this problem, a

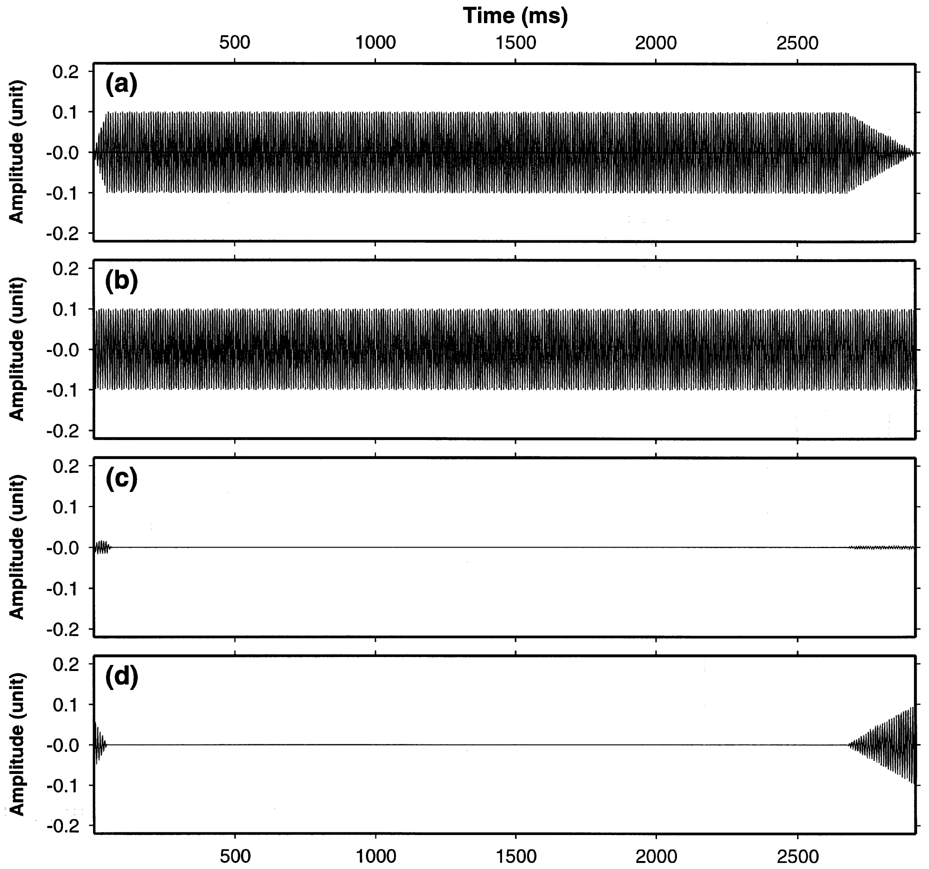



**Figure 7.** (a) Linear upsweep (10–60 Hz over 5000 ms) squeezed to 60 Hz with linear tapers at each end. (b) 60 Hz stationary sinusoidal reference sweep without any tapering, used in filter design to remove signal (a). (c) Resulting output from the notch algorithm method, and (d) from the optimum least-squares filtering algorithm method.

synthetic linear upsweep of 10 Hz to 60 Hz over 5000 ms with linear tapers of 500 ms on both ends is generated as a synthetic direct wave, which is subsequently squeezed into an undesired 60 Hz stationary sinusoidal signal by the squeeze (Fig. 7a). Note that the tapers on both ends still remain linear, but the one at the low-frequency end is shorter after squeezing than the other taper at the high-frequency end. If a reference sweep with no tapers (Fig. 7b) is used to design the OLS filter, a harmonic will be introduced into the filtered version, as shown in Fig. 7(d). This type of harmonic, caused by unmatched tapers, will degrade the performance of SFU so that a known pilot sweep is required if we wish to use the OLS filter. In cases where the pilot sweep specification is incomplete or not available (due, for example, to missing observer's logs), a linear taper should be assumed. This will minimize the harmonic distortion. However, the result obtained from the notch filter (Fig. 7c) shows that only slight

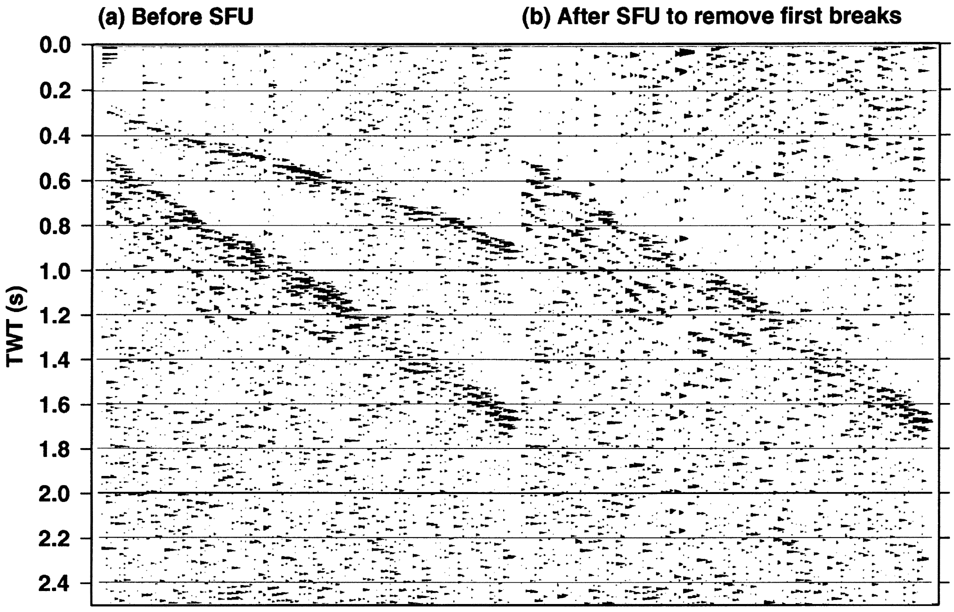



**Figure 8.** Correlated raw shot gather from Kola: linear sweep, 10–60 Hz over 20 s. (a) Before SFU and (b) after SFU to remove the first arrival. The same exponential gain is applied to both panels. Trace interval 50 m.

harmonic distortion occurs at each end, implying that SFU with the notch filter is insensitive to unmatched tapers in the filter design. In conclusion, we must use a known pilot sweep as the reference sweep in SFU with the OLS algorithm, whereas we do not need to worry about taper matching if we use a simple notch filter.

If the harmonic distortion is serious, we have to use SFU with the notch filter. In practice, we should use the OLS algorithm in SFU only when both the sweep tapers and the onset time of the direct wave (or other undesired signal) are known exactly.

## Field data examples

This section shows two field examples, with a linear and a non-linear sweep, respectively. The first example shows SFU with the stationary sinusoidal noise cancelling algorithm applied to a shot gather recorded with a linear sweep source, acquired in a crustal seismic reflection profile in the Kola Peninsula, Russia in 1992 (Smythe *et al.* 1994). The reflection data shown here are the 90 vertical component channels; the 180 other channels per shot, not shown here, recorded the two horizontal components. The source is a 20 s upsweep ranging from 10 Hz to 60 Hz, with 0.5 s linear tapers on each end. Both gathers are displayed with the same AGC window of 1 s for comparison purposes.

The unfiltered correlated shot gather is shown in Fig. 8(a). The 90-trace record is highly contaminated by mechanical sources near the Kola superdeep well, particularly

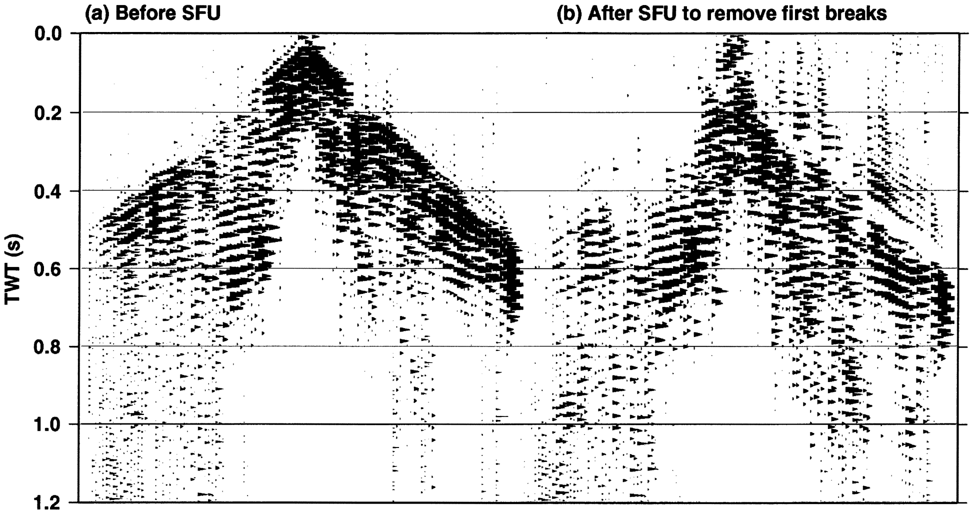



**Figure 9.** Correlated raw shot gather from Ohio: non-linear sweep, 20–120 Hz over 7 s. (a) Before SFU and (b) after SFU to remove the first arrival. The same exponential gain is applied to both panels. Trace interval 33 m.

on the channels to the right of the gather. The direct wave arrives between 0.3 and 0.9 s. The filtered output is shown in Fig. 8(b). Correlation analysis was used to pick automatically the onset time of the uncorrelated direct wave. SFU has done a good job of cancelling the first arrivals in the precorrelation domain, so that the amplitude of any correlation noise will be reduced to less than the amplitude of the genuine correlation signals from weak reflections. As mentioned above, the direct wave is removed without attenuating signal near the first arrivals. If, after the application of the filter, some trace of the first arrivals remains, it usually indicates that the onset time of the sweep is not determined precisely enough for SFU to work properly.

The second example is a split-spread shot gather using a non-linear sweep source, with frequency ranging from 20 Hz to 120 Hz over 7 s, acquired over the Serpent Mound structure in Ohio, USA. The direct wave is much stronger than that in the first example. Figure 9(a) is the correlated shot gather without application of SFU. Figure 9(b) shows the correlated outputs of SFU for the same shot gather as in Fig. 9(a). The filtered version has a much better signal-to-noise ratio, particularly in the shallow part of the section. This improvement is due to the fact that SFU suppresses the direct wave without having any adverse effect on the signal.

## Discussion

*The requirement in SFU filtering for uncorrelated data*

The recording of unsummed, uncorrelated vibroseis data was advocated over a decade ago by Stanley (1986), within the context of the reduction of ambient noise. This view met with some opposition from Scott (1987), who considered that to do so could actually degrade the field effort by depriving the field party observer of valuable quality control (QC) data. Stanley (1988) concurred that field stacking and correlation is an essential QC process, but amplified his earlier view that the output from the correlator-stacker should not be regarded as the primary raw data set. This discussion has been reproduced by Geyer (1989). However, digital data storage and processing costs are now so low that there is no good reason why unsummed, uncorrelated vibroseis field data should not be recorded as the archive field data set. The requirement for good field QC data can be met by stacking, field correlation and other preprocessing of the field data in the field recording truck, as required by the observer, using a parallel data stream. The essential point is that no irrevocable processing steps are taken. This recording method was adopted during the Kola three-component deep crustal vibroseis survey discussed above (Smythe *et al.* 1994).

Is it possible to recover the uncorrelated data trace from the correlated trace? The answer is probably yes (inasmuch as the uncorrelated trace can be estimated), if the pilot sweep is assumed, and if certain additional assumptions are made about the data. However, we have not investigated this; we only ask this question because it then follows that if the answer is affirmative, do we really need uncorrelated data after all? It is certainly obvious that vertical stacking (field summation) cannot be undone or reversed by any later processing, so that fact alone is a good argument for recording the unsummed data. We could therefore correlate in the field but store individual sweeps in unsummed but correlated form as the archive data set. This is the intermediate compromise that we adopted during the high-density trial 3D vibroseis survey of a potential radioactive waste repository (Smythe *et al.* 1995; Smythe 1996a,b).

We suspect that the assumptions required in estimating the uncorrelated trace from the correlated trace are such that the value of then applying any precorrelation process such as SFU is probably limited. In other words, any filter that can be applied to the synthesized or estimated uncorrelated trace (comprising a convolution of overlapping signal sweeps and noise) might just as well be applied to the correlated trace (comprising Klauder wavelets and correlation noise). However, this negative conclusion should be tested with some actual examples.

### *Application of SFU filtering to other unwanted signals*

We have demonstrated the usefulness of SFU filtering in removing first breaks. The method is equally applicable to the removal of other refracted events, unwanted shear waves, air blast, multiples, etc. The only requirements are that the precise time of the unwanted event be measurable, and that it is a sweep-generated event. The correlation side lobes from any event extend in time for the same period as the sweep length, on *each side* of the zero-lag peak. Thus a late event after correlation will have unwanted side lobes which may degrade data arriving before that event. Thus any large-amplitude unwanted event on correlated data should be removed by SFU filtering (in the precorrelation domain) to ensure that its side lobes are not degrading the data before or after the unwanted event itself.

## Conclusions

The 'squeeze-filter-unsqueeze' (SFU) algorithm is designed to remove any signal from uncorrelated vibroseis data which might be giving rise to correlation noise. We have demonstrated its efficacy at removing large-amplitude first breaks from synthetic and real data. The algorithm has the following costs and benefits:

- it requires uncorrelated data;
- it requires a knowledge of the source sweep;
- it is a single trace process (and no geometry information is required);
- it does not degrade the trace in any way;
- it has no special acquisition requirements;
- it works with a variety of source sweeps;
- it is cpu intensive, but automatic, so it is not labour intensive.

The version of SFU with the application of an OLS filter appears to be much more effective at cancelling the noise without attenuating signal frequency components near the terminal frequency than the version using a notch filter. However, if the tapers on either end of the sweep are unknown, severe harmonic distortion is introduced. In that case SFU with the notch filter should replace SFU with the OLS filter. The field data examples demonstrate that SFU requires the onset time of the direct-wave sweep to be known precisely. In cases where the onset time is known only approximately, a correlation analysis can be used to find the exact onset time needed for the squeezing and unsqueezing transforms to work automatically. If the auto-pick still fails or performs poorly, SFU with the notch filter should replace SFU with the OLS filter.

Even when the sweep formula is unknown or cannot be described as an explicit function of time, SFU is still valid for vibroseis data recorded with both linear or non-linear sweeps. In this case we search for and pick the zero nodes on the recorded sweep directly, without deriving an explicit transformation algorithm.

Although we have investigated the SFU algorithm as a means of removing correlation noise from first breaks, there is no reason why SFU should not be applied to the removal of any undesired signal, as long as the exact onset time of the unwanted signal in the precorrelation domain is known or determinable.

## References

Bernhardt T. and Peacock J.H. 1978. Encoding techniques for the vibroseis system. *Geophysical Prospecting* **26**, 184–193.

Cunningham A.B. 1979. Some alternate vibrator signals. *Geophysics* **44**, 1901–1921.

Edelmann H.A.K. 1966. New filtering methods with Vibroseis. *Geophysical Prospecting* **14**, 455–469.

Edelmann H.A.K. and Werner H. 1982. Combined sweep signals for correlation noise suppression. *Geophysical Prospecting* **30**, 786–813.

Ensing L. 1983. The autobalancer – an automatic system for the reduction of powerline interference with seismic signals. *Geophysical Prospecting* **31**, 591–607.

Geyer R.L. (ed.) 1989. *Vibroseis*. Geophysics Reprint Series No. 11. Society of Exploration Geophysicists, Tulsa, OK.

Klauder J.K., Price A.C., Darlington S. and Albersheim W.J. 1960. The theory and design of chirp-radars. *The Bell System Technical Journal* **39**, 745–808.

Linville A.F. and Meek R.A. 1992. Cancelling stationary sinusoidal noise. *Geophysics* **57**, 1493–1501.

Press W.H., Flannery B.P., Teukolsky S.A. and Vetterling W.T. 1989. Fourier transform spectral methods. In: *Numerical Recipes*, pp. 381–449. Cambridge University Press.

Scott P.J. 1987. Ambient noise in Vibroseis surveys. *First Break* **5**, 310.

Smythe D.K. 1996a. 3-D seismic reflection trial: data acquisition report. UK Nirex Ltd Report No. 622. Harwell, UK.

Smythe D.K. 1996b. 3-D seismic reflection trial survey of the potential nuclear waste repository zone, Sellafield, West Cumbria. UK Nirex Ltd Report No. 760. Harwell, UK.

Smythe D.K., Holt J.M., Elstob M. and Robson C. 1995. A high-resolution 3D vibroseis trial survey of a potential nuclear waste repository. 57th EAGE meeting, Glasgow, UK, Extended Abstracts, C054.

Smythe D.K., Smithson S.B., Gillen C., Humphreys C., Kristoffersen Y., Karaev N.A., Garipov V.Z., Pavlenkova N.I. and the Kola-92 Working Group. 1994. Project images crust, collects seismic data in world's largest borehole. *EOS Transactions of the American Geophysical Union* **75**, 473–476.

Stanley P.J. 1986. Some comments regarding ambient noise in Vibroseis surveys. *First Break* **4**, 31–33.

Stanley P.J. 1988. A neglected area of Vibroseis development. *First Break* **6**, 113–116.

Werner H. and Krey T. 1979. Combisweep – a contribution to sweep techniques. *Geophysical Prospecting* **27**, 78–106.

Wiener N. 1949. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. M.I.T. Press and John Wiley & Sons, Inc.

Yilmaz O. 1987. *Seismic Data Processing*, pp. 501–506. Society of Exploration Geophysicists, Tulsa, OK.