# Data regularization and redatuming using Newton's method

*Robert J. Ferguson*, and Sergey B. Fomel, University of Texas at Austin*

## SUMMARY

We present a Newton-method approach for regularizing seismic data simultaneous with wavefield extrapolation. Our approach is based on $\omega - x$ approximations to forward extrapolation, and to the composition of back extrapolation with the adjoint of back extrapolation. The advantage of this method is that it is robust in the presence of highly irregular trace spacing, and fast computation is achieved.

We demonstrate our approach using synthetic data derived by finite differences that represent propagation of a line source and a number of point sources a large distance through a medium with a strong lateral velocity contrast. The line source is used to generate a simple, linear in appearance structure in the data, and the point sources are used to generate a set of steep, conflicting dips. After synthesizing the data, we set to zero traces selected randomly such that a highly irregular recording geometry is simulated. We attempt to recover the line and point sources by back propagating the irregular data through the medium using an expensive, but accurate, generalized regularizing/extrapolation method based on Newton's method. We compare the resulting estimate of the sources to those recovered using generalized PSPI. Based on visual inspection, we find that the Newton method regularizes the data, and returns a more accurate estimate of the sources.

We then recover the sources using a fast approximation to the Newton method that is based on the $\omega - x$ approximation. We find that this approximation recovers an estimate of the sources that compares favorably with the non-approximation Newton method for less than a $10^{th}$ the computational cost for 2D data. For 3D, we expect cost to be a $100^{th}$ the cost of the non-approximated method. We compare this result also to simple back-propagation using plain $\omega - x$, and we find this result inferior to our Newton method approach.

We feel our Newton method may have great utility in statics correction for 3D data acquired with irregular geometry onshore, and for redatuming offshore where OBS/OBC and towed streamer geometries are irregular.

## WAVEFIELD EXTRAPOLATION BY LEAST SQUARES

For mixed determined inverse problems like wavefield extrapolation, weighted, damped-least squares is a useful way to cope with incomplete, noisy data (Menke, 1989; Tarantola, 1987). When wavefield extrapolation operators are linear, conjugate gradient methods and the Newton method are efficient approaches to compute the extrapolated wavefield; conjugate gradient methods converge in a finite number of iterations, and the Newton method converges in a single iteration (Tarantola, 1987, pg. 250).

For linear operators based on PSPI-like (Gazdag and Sguazzero, 1984) approximations (Kuehl and Sacchi, 2002), conjugate gradient methods are perhaps most suitable because the adjoint operator is known (Margrave and Ferguson, 1999), and a reliable estimate of $\psi_{z+\Delta z}$ is obtainable after a few iterations (Kuehl and Sacchi, 2002).

Efficiency in the PSPI approach applied to conjugate-gradient methods is seen in computation of the direction of steepest accent. As part of each iteration $j$, the direction of steepest ascent $\kappa_j$ is computed (Tarantola, 1987, pg. 251)

$$\kappa_j = \left[\varepsilon^2 W_m U^A_{-\Delta z} W_e \left[U_{\Delta z}\psi_{z+\Delta z} - \psi_z\right] + \psi_z\right], \tag{1}$$

where $W_e$ and $W_m$ are a weighting operator and a smoothing operator respectively, and $\varepsilon^2$ is a scalar that controls the amount of smoothing (Menke, 1989, pg. 53 - 54). Computation of $\kappa$ dominates computational cost as the solution converges to the optimal extrapolated wavefield $\psi_{z+\Delta z}$.

Operators $U_{\Delta z}$ and $U^A_{-\Delta z}$, when approximated by PSPI-like algorithms, have cost $\propto MN\log MN$, where $M$ and $N$ are the in-line and cross-line dimensions. Then, because $W_e$ is usually diagonal, and because $\varepsilon^2 W_m$ is usually tridiagonal, $U_{\Delta z}$ and $U^A_{-\Delta z}$ contribute the bulk of the computational cost of $\kappa_j$ (equation (1)), so actual cost is $\propto JMN\log MN$, where $J$ is the number of iterations in the conjugate gradient solution.

For the Newton method, operator $U^A_{-\Delta z} W_e U_{-\Delta z}$ must be calculated, and $\psi_{z+\Delta z}$ computed as follows (Tarantola, 1987, pgs. 251)

$$\psi_{z+\Delta z} = \left[U^A_{-\Delta z} W_e U_{-\Delta z} + \varepsilon^2 W_m\right]^{-1} U^A_{-\Delta z} W_e \psi_z. \tag{2}$$

To avoid the required matrix multiplications in the denominator as well as the numerator, efficient approximations to $U^A_{-\Delta z} W_e U_{-\Delta z}$ and $U^A_{-\Delta z}$ are desirable. Then, for example, $LU$ factorization is used to compute $\psi_{z+\Delta z}$ as

$$\psi_{z+\Delta z} = T_U^{-1} T_L^{-1} b_{z+\Delta z}, \tag{3}$$

where $T_U$ and $T_L$ are upper and lower triangular matrices that correspond to the $LU$ factorization of $U^A_{-\Delta z} W_e U_{-\Delta z} + \varepsilon^2 W_m$, and

$$b_{z+\Delta z} = \tilde{U}^A_{-\Delta z} W_e \psi_z \tag{4}$$

is computed by scaling wavefield $\psi_z$ with diagonal-weighing operator $W_e$, and then extrapolating the result with efficient, approximate extrapolator $\tilde{U}$.

In the two following sections, we introduce Fourier integral expressions for $U_{\Delta z}$ and adjoint $U^A_{\Delta z}$, and then we develop an exact, integral representation for composition $U^A_{-\Delta z} W_e U_{-\Delta z}$ that resides in the denominator of equation (2). We find that this composition is extremely expensive to compute, so we approximate the extrapolators within the integrals as series, and then truncate to a manageable number of terms.

For eight terms, our truncated least-squares operator implemented using the Newton method is accurate for dips up to 65 - 80 degrees (depending on velocity), and for 2D data, our eight-term approximation reduces the computational cost of the Newton method by a factor of $40/N$, where $N$ is the number of traces. Accuracy for greater angles is achieved simply by increasing the number of terms.

## EXTRAPOLATOR $U$ AND ADJOINT $U^A$

For monochromatic wavefields, extrapolated wavefield $\psi_{z+\Delta z}$ is given by a general expression for one-way extrapolation through a thin slab of thickness $\Delta z$ as

$$\psi_{z+\Delta z} = U_{\Delta z}\psi_z \tag{5}$$

where $U_{\Delta z}\psi_z$ is an integral operator given by

$$\left[U_{\Delta z}\psi_z(y)\right](x) = \frac{1}{(2\pi)^2}\int e^{-i\langle k_x, x-y\rangle}\alpha_{\Delta z}(x,k_x)\,\psi_z(y)\,dy\,dk_x \tag{6}$$

(Margrave and Ferguson, 1999; Rousseau and de Hoop, 2001).

Variable $y = \{y_1, y_2\}$ represents the space coordinates of input wavefield $\psi_z$ at depth $z$, and variables $x = \{x_1, x_2\}$ and $k_x = \left\{k_{x_1}, k_{x_2}\right\}$ are, respectively, space and wavenumber coordinates of output wavefield $\psi_{z+\Delta z}$ at depth $z+\Delta z$. Limits of integration $\pm\infty$ are omitted for brevity, and compact notation $\langle a,b\rangle$ is equivalent to dot product $a_1 b_1 + a_2 b_2$. Extrapolator $\alpha$ is

$$\alpha_{\Delta z} = e^{i\Delta z k_z}, \tag{7}$$

and, for frequency $|\omega| = \omega$, wavenumbers $k_z$ are

$$k_z(x,k_x) = \begin{cases} \operatorname{sgn}(\Delta z)\, k(x) \sqrt{1 - \frac{|k_x|^2}{k^2(x)}} & \text{if} \quad \frac{|k_x|^2}{k^2(x)} \le 1; \\ i \operatorname{sgn}(\Delta z)\, k(x) \sqrt{1 - \frac{|k_x|^2}{k^2(x)}} & \text{if} \quad \frac{|k_x|^2}{k^2(x)} > 1, \end{cases} \tag{8}$$

where

$$k(x,\omega) = \frac{\omega}{v(x)}, \tag{9}$$

and $v$ is seismic velocity that varies laterally within the thin slab.

Adjoint $U^A$ applied to wavefield $\psi_z$ reverses the direction of extrapolation of $\psi_z$, and back-propagated wavefield $\psi_{z-\Delta z}$ is computed as

$$\psi_{z-\Delta z} = U^A_{\Delta z}\, \psi_z, \tag{10}$$

where $U^A_{\Delta z}\, \psi_z$ is an integral operator given by

$$\left[U^A_{\Delta z}\, \psi_z(y)\right](x) = \frac{1}{(2\pi)^2} \int e^{-i\langle k_x, x-y\rangle} \alpha^*(y,k_x,\Delta z)\, \psi_z(y)\, dk_x\, dy \tag{11}$$

(Margrave and Ferguson, 1999), where $\alpha^*(\Delta z) = \alpha(-\Delta z)$ is the complex conjugate of $\alpha$.

For the Newton method (equation (2)), equation (11) provides an analytic expression for computing $U^A_{-\Delta z}$.

**Composition operator $U^A W_e U$**
Using space variables $y'$ and wavenumbers $k'_x$ to represent the lateral-coordinate pairs at depth $z-\Delta z$, apply weighting $W_e$ and back propagation $U_{-\Delta z}$ (equation (6) parameterized with $-\Delta z$) to arbitrary wavefield $\psi_z$

$$[W_e U_{-\Delta z}\, \psi_z(y)](y') = \frac{1}{(2\pi)^2} \int e^{-i\langle k'_x, y'-y\rangle} \tilde{\alpha}(y', k'_x, -\Delta z)\, \psi_z(y)\, dy\, dk'_x, \tag{12}$$

where $\tilde{\alpha}$ is the result of diagonal-weighting operator $W_e$ applied to extrapolator $\alpha$

$$\tilde{\alpha}(y') = \int \alpha(y'')\, W_e(y'-y'')\, dy'' = W_e(y')\, \alpha(y'). \tag{13}$$

Adjoint extrapolation $U^A_{-\Delta z}$ (equation (11) parameterized with $-\Delta z$) applied to equation (12) is

$$\begin{aligned} &\left[U^A_{-\Delta z} W_e U_{-\Delta z}\, \psi_z(y)\right](x) = \\ &\frac{1}{(2\pi)^4} \int e^{-i\langle k_x, x-y'\rangle} e^{-i\langle k'_x, y'-y\rangle} \\ &\alpha^*(y', k_x, -\Delta z)\, \tilde{\alpha}(y', k'_x, -\Delta z)\, \psi_z(y)\, dy\, dk'_x\, dy'\, dk_x. \end{aligned} \tag{14}$$

As it is, equation (14) is costly to compute. Though integral $\int_y$ is simply the FFT of wavefield $\psi_z$, symmetry in the remaining Fourier kernals is broken in general by $\alpha^* \tilde{\alpha}$. For 2D data, integrals $\int_{y'}$, $\int_{k_x}$, and $\int_{k'_x}$, therefore, may not be computed with FFTs, and without approximation, cost for each integral is $\propto N^2$ where $N$ is the number of traces. For 3D data, this cost is quadratic.

Together with equation (11), equation (14) provides the analytic framework to implement the Newton method given in equation (2).

## APPROXIMATION

Because the Newton method implemented using equation (14) is costly to implement, some kind of approximation must be considered. A common expansion of $\alpha$ as given by equation (7) is (Berkhout, 1985, pg. 439)

$$\alpha(y,k_x,\Delta z) = e^{i\Delta z k(y)} \sum_{p=0}^{\infty} \gamma_p(y,\Delta z)\, k_x^{2p}, \tag{15}$$

where $k(y) = \omega / v(y)$, $\gamma_0 = 1$, and $\gamma_1$, $\gamma_2$, and $\gamma_3$ are given in Table 1. This exapansion is known, generaly, as the $\omega - x$ approximation, and when implemented in space coordinates, it is known as Fourier finite-difference. Depending on the maximum value of integer $p$ (equation (15)), a maximum value for geological dip is implied and, for example, for $\mathrm{MAX}(p) = 4$, an operator suitible for dip between 60 and 80 degrees is approximated (Berkhout, 1985, pg. 344).

From equation (14), $\alpha^*(-\Delta z)\, \tilde{\alpha}(-\Delta z)$ becomes

$$\begin{aligned} &\alpha^*(y', k_x, -\Delta z)\, \tilde{\alpha}(y', k'_x, -\Delta z) = \\ &\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \gamma^*_p(y', -\Delta z)\, \tilde{\gamma}_q(y', -\Delta z)\, (k_x)^{2p} \left(k'_x\right)^{2q}, \end{aligned} \tag{16}$$

where, from equation (13),

$$\tilde{\gamma}_q(y') = W(y')\, \gamma_q(y'). \tag{17}$$

The $*$ on $\gamma^*_p$ indicates complex conjugate. Substitute equation (16) for $\alpha^* \tilde{\alpha}$ in equation (14) to get

$$\begin{aligned} &\left[U^A_{-\Delta z} W_e U_{-\Delta z}\, \psi_z(x)\right](x) = \\ &\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} i^{-2(p+q)} (\partial_x)^{2p} \left[\gamma^*_p(x,-\Delta z)\, \tilde{\gamma}_q(x,-\Delta z)\, (\partial_x)^{2q}\, \psi_z(x)\right]. \end{aligned} \tag{18}$$

In arriving at equation (18) from equations (16) and (14), we eliminate integration over $y$, $y'$, $k_x$, and $k'_x$ using $FFT_{y\to k'_x}\{\psi_z\}$ to compute integral $\int_y$, and we use the $\partial_x \Leftrightarrow i k_x$ relationship to convert $\int_{k_x}$ and $\int_{k^{\partial}_x}$ to spatial derivatives. The $\int_{y'}$ integral convertes $y'$ to output coordinate $x$.

Without truncation, equation (18) is an exact representation of $U_{-\Delta z} W_e U_{-\Delta z}\, \psi_z$ where an infinite series of spatial derivatives of $\psi_z$, $\gamma^*$, and $\tilde{\gamma}$ is summed.

Computational efficiency using equation (18) is achieved when $0 \le p << \infty$ and $0 \le q << \infty$ result in a small number of terms. For example, for $0 \le p \le 4$ and $0 \le q \le 4$, equation (18) is a series of eight terms. The $8^{th}$ term in this series corresponds to a set of $2^{nd}$, $4^{th}$, $6^{th}$, and $8^{th}$ order partial derivatives applied to wavefield $\psi_z$ and velocity $v$. These derivatives in term eight are toeplitz matricies with two, four, six, and eight non-zero diagonals respectively and a combined computational cost dominated by the $8^{th}$ order derivative. Cost for terms seven, six, and five are likewise dominated by the $8^{th}$ order derivative, with terms four, three, and two dominated by $6^{th}$, $4^{th}$, and $2^{nd}$ order derivatives respectively. Term one is $W_e$ and is essentially free. In 2D, therefore, we estimate the total cost of the eight-term operator to be $\propto 40N$ for $N$ traces.

Our approximation, therefore, is at an order of magnitude less expensive to implement in 2D, and for 3D data, we expect at least a second order decrease in cost.

## EXAMPLE

A test velocity model and corresponding seismic data are given in Figures 1a and b. The velocity is a step function that varies abruptly between 2000 m/s on the left to 3000 m/s on the right. Embedded in this medium are a line source and a number of point sources. Upon excitation of the sources, energy is propagated through the medium with a 9-point, finite-difference algorithm into an irregular receiver array 200 m above. The design of the model, receiver array, and source ensures a recorded wavefield that has a simple event due to the line source, plus a set of conflicting dips due to the point sources, plus the added complexity of irregular sampling. The large depth interval is intended to reveal instability intrinsic to the operators we will compare.

For accurate redatuming and regularization, reversing the wavefield (Figure 1b) through this model (Figure 1a) should recover the input sources, and fill in missing amplitudes; the linear event should be continuous, and the point sources should be focused. As can be seen in Figure 2a, the least-squares Newton method based on equation (14) recovers a coherent estimate of the line source, and the point sources are well focused.

In Figure 2b, the source wavefield is estimated more simply by reversing the recorded wavefield using plain PSPI. Comparing Figures 2a and 2b shows that, besides returning a focused, regularized result, the Newton method reduces significantly the imprint of the extrapolation operator through *migration deconvolution* (Jianxing et al., 2001).

Figure 3a gives the result from regularization and redatuming according to the approximate, least-squares operator based on equation (18). This result, though not quite as coherent in appearance as 2a, is still a good estimate of the line source and point sources, with a minimized operator imprint. The result based on simple reversal of the $\omega - x$ operator is incoherent (Figure 3b), and it suffers from operator imprint everywhere, and from operator instability to the right side of the velocity step.

## CONCLUSIONS

We present in this abstract a regularization/redatuming approach that is based on the Newton method. Using synthetic data derived using finite differences, we demonstrate that our method is robust in the presence of highly irregular sampling. We also demonstrate the reduction in operator imprint that is a partial result of our Newton method approach when extrapolation distances are large, and velocity contrast is strong.

Because the Newton method requires the computation of a composition of operators with a large associated cost, we present a efficient approximation. We find that for an operator equivalent to a 65 - 80 degree $\omega - x$ operator, we reduce computational cost by 10 in 2D. We speculate that cost in 3D will be reduced by 100. Using our approximate operator, we find that most of the the desirable aspects of data regularization, and of reducing operator imprint are preserved relative to the full technique.

In areas where surface limitations restrict our ability to acquire seismic data with regularly spaced grids, and in areas where extreme topography impact the coherence of data, we see great utility in this efficient method that solves these combined effects. In future, we intend to study alternative approximations to the full inversion, and compare directly the Newton approach to alternative approaches such as conjugate gradients.

## ACKNOWLEDGMENTS

The first author thanks the sponsors of the Exploration Development Geophysics Education and Research (EDGER) Forum, University of Texas, Austin, for suporting this research.

## REFERENCES

Berkhout, A. J., 1985, Seismic migration. imaging of acoustic energy by wavefield extrapolation. a. theoretical aspects: Elsevier.

Gazdag, J., and Sguazzero, P., 1984, Migration of seismic data by phase-shift plus interpolation: Geophysics, **49**, no. 02, 124–131.

Jianxing, H., Schuster, G. T., and Valasek, P. A., 2001, Poststack migration deconvolution: Geophysics, **66**, no. 3, 939 – 952.

| $j$ | $\gamma_j$ |
|---|---|
| 0 | 1 |
| 1 | $-\frac{i\Delta z}{2k}$ |
| 2 | $\frac{-1/8\,i\Delta z k - 1/8\,\Delta z^2 k^2}{k^4}$ |
| 3 | $\frac{-1/16\,i\Delta z k - 1/16\,\Delta z^2 k^2 + 1/48\,i\Delta z^3 k^3}{k^6}$ |
| 4 | $\frac{-5/128\,i\Delta z k - 5/128\,\Delta z^2 k^2 + 1/64\,i\Delta z^3 k^3 + 1/384\,\Delta z^4 k^4}{k^8}$ |

Table 1: Table of $\omega - x$ variables.

Kuehl, H., and Sacchi, M., 2002, Robust AVP estimation using least-squares wave-equation migration: 71st Ann. Internat. Mtg, 281–284.

Margrave, G. F., and Ferguson, R. J., 1999, Wavefield extrapolation by nonstationary phase shift: Geophysics, **64**, no. 4, 1067–1078.

Menke, W. ., 1989, Geophysical data analysis: Discrete inverse theory: Academic Press.

Rousseau, J. H. L., and de Hoop, M. V., 2001, Modeling and imaging with the scalar generalized-screen algorithms in isotropic media: Geophysics, **66**, no. 5, 1551–1568.

Tarantola, A. ., 1987, Inverse problem theory: Elsevier Science.

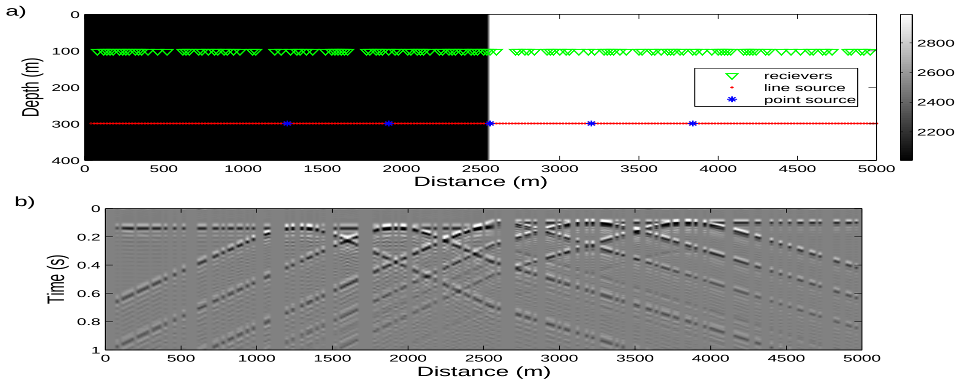


Figure 1: Step-function velocity and corresponding seismic data (synthetic). a) Velocity varies from a low value on the left to a high value on the right abruptly at 2500 m distance. The medium is 200 m thick. A line source (minimum phase) is represented by the line at 300 m depth. Receivers at 100 m depth are spaced randomly with a minimum spacing of 20 m. b) Synthetic seismic data.

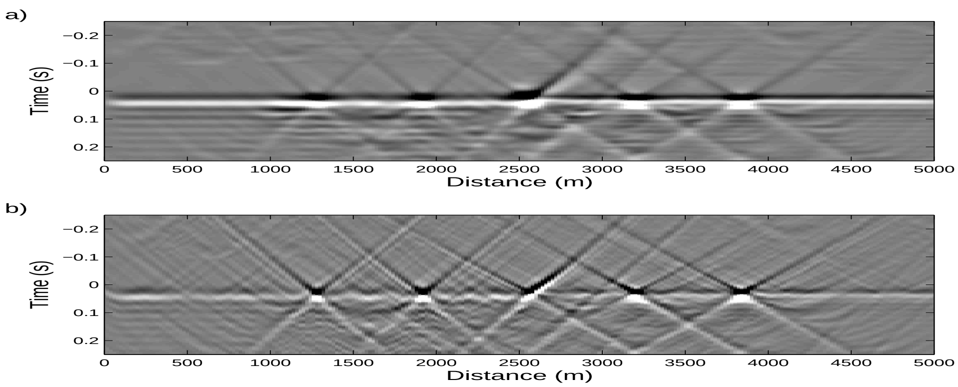


Figure 2: Regularization and redatuming of the data from Figure 1b) by least-squares using generalized PSPI (a), and by generalized PSPI (non least-squares) (b). Least-squares PSPI returns the most coherent estimate of the line and point sources. .

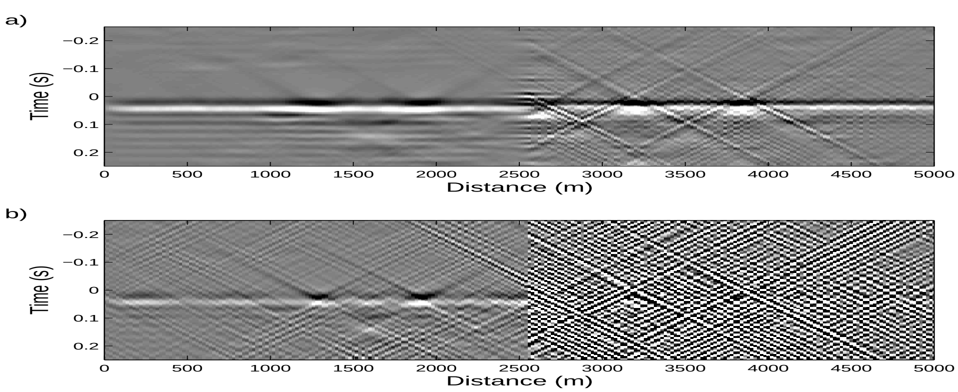


Figure 3: Regularization and redatuming of the data from Figure 1b) by least-squares using the $\omega - x$ based approximation (a), and redatuming using standard (non least-squares) $\omega - x$ (b). Least-squares $\omega - x$ returns a more coherent estimate of the line and point sources compared to standard $\omega - x$. For this model, $\omega - x$ operators are very unstable as shown in (b). Least-squares $\omega - x$, however, controls the stability, and the result compares favorably with least-squares PSPI but at $\sim 1/10$ the cost.