Chapter 11

Continuous Inverse Theory and Tomography

11.1 THE BACKUS-GILBERT INVERSE PROBLEM

While continuous inverse problems are not the main subject of this book, we will cover them briefly to illustrate their relationship to discrete problems. Discrete and continuous inverse problems differ in their assumptions about the model parameters. Whereas the model parameters are treated as a finite-length vector in discrete inverse theory, they are treated as a continuous function in continuous inverse theory. The standard form of the continuous inverse problem is

$$d_i = \int_a^b G_i(z) m(z) \mathrm{d}z \tag{11.1}$$

when the model function $m(z)$ varies only with one parameter, such as depth z. When the model function depends on several—say L—variables, then Equation (11.1) must be generalized to

$$d_i = \int_V G_i(\mathbf{x}) m(\mathbf{x}) \mathrm{d}^L x \tag{11.2}$$

where $\mathrm{d}^L x$ is the volume element in the space of $\mathbf{x}$.

The "solution" of a discrete problem can be viewed as either an estimate of the model parameter vector $\mathbf{m}^{\text{est}}$ or a series of weighted averages of the model parameters, $\mathbf{m}^{\text{avg}} = \mathbf{R}\mathbf{m}^{\text{true}}$, where $\mathbf{R}$ is the resolution matrix (see Section 4.3). If the discrete inverse problem is very underdetermined, then the interpretation of the solution in terms of weighted averages is most sensible, since a single model parameter is very poorly resolved. Continuous inverse problems can be viewed as the limit of discrete inverse problems as the number of model parameters becomes infinite, and they are inherently underdetermined. Attempts to estimate the model function $m(\mathbf{x})$ at a specific point $\mathbf{x} = \mathbf{x}'$ are futile. All determinations

Geophysical Data Analysis: Discrete Inverse Theory. DOI: 10.1016/B978-0-12-397160-9.00011-4

of the model function must be made in terms of local averages, which are simple generalizations of the discrete case, $m_i^{\text{avg}} = \sum_j G_{ij}^{-g} d_j = \sum_j R_{ij} m_j^{\text{true}}$ where $R_{ij} = \sum_k G_{ik}^{-g} G_{kj}$

$$m^{\text{avg}}(\mathbf{x}') = \sum_{i=1}^{N} G_i^{-g}(\mathbf{x}') d_i = \int R(\mathbf{x}', \mathbf{x}) m^{\text{true}}(\mathbf{x}) \mathrm{d}^L x \tag{11.3}$$

where

$$R(\mathbf{x}', \mathbf{x}) = \sum_{i=1}^{N} G_i^{-g}(\mathbf{x}') G_i(\mathbf{x}) \tag{11.4}$$

Here, $G_i^{-g}(\mathbf{x}')$ is the continuous analogy to the generalized inverse G_{ij}^{-g} and the averaging function $R(\mathbf{x}', \mathbf{x})$ (often called the *resolving kernel*) is the analogy to the model resolution matrix R_{ij}. The average is localized near the target point $\mathbf{x}'$ if the resolving kernel is peaked near $\mathbf{x}'$. The solution of the continuous inverse problem involves constructing the most peaked resolving kernel possible with a given set of measurements, that is, with a given set of data kernels, $G_i(\mathbf{x})$. The spread of the resolution function is quantified by (compare with Equation (4.23))

$$J(\mathbf{x}') = \int w(\mathbf{x}', \mathbf{x}) R^2(\mathbf{x}', \mathbf{x}) \mathrm{d}^L x \tag{11.5}$$

Here, $w(\mathbf{x}', \mathbf{x})$ is a nonnegative function that is zero at the point $\mathbf{x}'$ and that grows monotonically away from that point. One commonly used choice is the quadratic function $w(\mathbf{x}', \mathbf{x}) = |\mathbf{x}' - \mathbf{x}|^2$. Other, more complicated functions can be meaningful if the elements of $\mathbf{x}$ have an interpretation other than spatial position. After inserting the definition of the resolving kernel (Equation (11.3)) into the definition of the spread (Equation (11.5)), we find

$$\begin{aligned} J(\mathbf{x}') &= \int w(\mathbf{x}', \mathbf{x}) R(\mathbf{x}', \mathbf{x}) R(\mathbf{x}', \mathbf{x}) \mathrm{d}^L x \\ &= \int w(\mathbf{x}', \mathbf{x}) \sum_{i=1}^{N} G_i^{-g}(\mathbf{x}') G_i(\mathbf{x}) \sum_{j=1}^{N} G_j^{-g}(\mathbf{x}') G_j(\mathbf{x}) \mathrm{d}^L x \\ &= \sum_{i=1}^{N} \sum_{j=1}^{N} G_i^{-g}(\mathbf{x}') G_j^{-g}(\mathbf{x}') \int w(\mathbf{x}', \mathbf{x}) G_i(\mathbf{x}) G_j(\mathbf{x}) \mathrm{d}^L x \\ &= \sum_{i=1}^{N} \sum_{j=1}^{N} G_i^{-g}(\mathbf{x}') G_j^{-g}(\mathbf{x}') [\mathbf{S}(\mathbf{x}')]_{ij} \end{aligned} \tag{11.6}$$

where

$$[\mathbf{S}(\mathbf{x}')]_{ij} = \int w(\mathbf{x}', \mathbf{x}) G_i(\mathbf{x}) G_j(\mathbf{x}) \mathrm{d}^L x \tag{11.7}$$

Equation (11.7) might be termed an *overlap integral*, since it is large only when the two data kernels overlap, that is, when they are simultaneously large in the same region of space. The continuous spread function has now been

manipulated into a form completely analogous to the discrete spread function in Equation (4.26). The generalized inverse that minimizes the spread of the resolution is the precise analogy of Equation (4.34)

$$G_l^{-g}(\mathbf{x}') = \frac{\sum_{i=1}^{N} u_i [\mathbf{S}^{-1}(\mathbf{x}')]_{il}}{\sum_{i=1}^{N}\sum_{j=1}^{N} u_i [\mathbf{S}^{-1}(\mathbf{x}')]_{ij} u_j} \quad \text{where} \quad u_i = \int G_i(\mathbf{x}) \mathrm{d}^L x \tag{11.8}$$

11.2 RESOLUTION AND VARIANCE TRADE-OFF

Since the data $\mathbf{d}$ are determined only up to some error quantified by the covariance matrix [cov $\mathbf{d}$], the localized average $m^{\text{avg}}(\mathbf{x}')$ is determined up to some corresponding error

$$\text{var}[m^{\text{avg}}(\mathbf{x}')] = \sum_{i=1}^{N}\sum_{j=1}^{N} G_i^{-g}(\mathbf{x}')[\text{cov } \mathbf{d}]_{ij} G_j^{-g}(\mathbf{x}') \tag{11.9}$$

As in the discrete case, the generalized inverse that minimizes the spread of resolution may lead to a localized average with large error bounds. A slightly less localized average may be desirable because it may have much less error. This generalized inverse may be found by minimizing a weighted average of the spread of resolution and size of variance

$$\text{minimize} J'(\mathbf{x}') = \alpha \int w(\mathbf{x}', \mathbf{x}) R^2(\mathbf{x}', \mathbf{x}) \mathrm{d}^L x + (1-\alpha)\text{var}[m^{\text{avg}}(\mathbf{x}')] \tag{11.10}$$

The parameter α, which varies between 0 and 1, quantifies the relative weight given to spread of resolution and size of variance. As in the discrete case (see Section 4.10), the corresponding generalized inverse can be found by using Equation (11.7), where all instances of $\mathbf{S}(\mathbf{x}')$ are replaced by

$$[\mathbf{S}'(\mathbf{x}')]_{ij} = \alpha \int w(\mathbf{x}', \mathbf{x}) G_i(\mathbf{x}) G_j(\mathbf{x}) \mathrm{d}^L x + (1-\alpha)[\text{cov}\, \mathbf{d}]_{ij} \tag{11.11}$$

Backus and Gilbert (1968) prove a number of important properties of the trade-off curve (Figure 11.1), which is a plot of size of variance against spread of resolution, including the fact that the variance decreases monotonically with spread.

11.3 APPROXIMATING CONTINUOUS INVERSE PROBLEMS AS DISCRETE PROBLEMS

A continuous inverse problem can be converted into a discrete one with the assumption that the model function can be represented by a finite number M of coefficients, that is

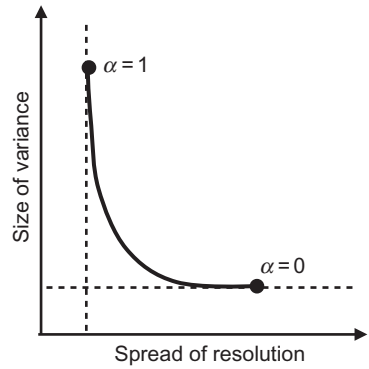

FIGURE 11.1 Typical trade-off of resolution and variance for a linear continuous inverse problem. Note that the size(spread) function decreases monotonically with spread and that it is tangent to the two asymptotes at the endpoints $\alpha = 1$ and $\alpha = 0$.

$$m(\mathbf{x}) \approx \sum_{j=1}^{N} m_j f_j(\mathbf{x}) \tag{11.12}$$

The particular choice of the functions $f_j(\mathbf{x})$ implies particular *a priori* information about the behavior of $m(\mathbf{x})$, so the solution one obtains is sensitive to the choice. One commonly used set of functions assumes that the model is constant within certain subregions V_j of the space of model parameters (e.g., voxels). In this case, $f_j(\mathbf{x})$ is unity inside V_j and zero outside it, and m_j is the value of the model in each voxel. In one dimension, where the x-axis is divided into intervals of width Δx

$$m(x) \approx \sum_{j=1}^{M} m_j f_j(x) \quad \text{with} \quad f_j(x) = H(x - j\Delta x) - H(x - (j+1)\Delta x) \tag{11.13}$$

Here $H(x - \xi)$ is the Heaviside step function, which is zero for $x < \xi$ and unity for $x > \xi$. Many other choices of $f_j(\mathbf{x})$ are encountered, based on polynomial approximations, splines, and truncated Fourier series representations of $m(\mathbf{x})$.

Inserting Equation (11.12) into the standard form of the continuous problem (Equation 11.2) leads to a discrete problem

$$d_i = \int G_i(\mathbf{x}) \sum_{j=1}^{m} m_j f_j(\mathbf{x}) \mathrm{d}^L x = \sum_{j=1}^{M} \left\{ \int G_i(\mathbf{x}) f_j(\mathbf{x}) \mathrm{d}^L x \right\} m_j \tag{11.14}$$

So the discrete form of the equation is $\mathbf{d} = \mathbf{G}\mathbf{m}$

$$d_i = \sum_{j=1}^{M} G_{ij} m_j \quad \text{with} \quad G_{ij} = \int G_i(\mathbf{x}) f_j(\mathbf{x}) \mathrm{d}^L x \tag{11.15}$$

In the special case of voxels of volume V_j centered on the points $\mathbf{x}^{(j)}$, Equation (11.15) becomes

$$G_{ij} = \int_{V_j} G_i(\mathbf{x}) \mathrm{d}^L x \tag{11.16}$$

If the data kernel varies slowly with position so that it is approximately constant in the voxel, then we may approximate the integral as just its integrand evaluated at $\mathbf{x}^{(j)}$ times the volume V_j:

$$G_{ij} = \int_{V_j} G_i(\mathbf{x}) \mathrm{d}^L x \approx G_i(\mathbf{x}^{(j)}) V_j \tag{11.17}$$

In one dimension, this is equivalent to the Reimann summation approximation to the integral

$$d_i = \int G_i(x) m(x) \mathrm{d}x \approx \sum_{j=1}^{M} \{G_i(j\Delta x)\Delta x\} m_j = \sum_{j=1}^{M} G_{ij} m_j \tag{11.18}$$

Two different factors control the choice of the size of the voxels. The first is dictated by an *a priori* assumption of the smoothness of the model. The second becomes important only when one uses Equation (11.17) in preference to Equation (11.16). Then the subregion must be small enough that the data kernels $G_i(\mathbf{x})$ are approximately constant in the voxel. This second requirement often forces the voxels to be much smaller than dictated by the first requirement, so Equation (11.16) should be used whenever the data kernels can be integrated analytically. Equation (11.17) fails completely whenever a data kernel has an integrable singularity within the subregion. This case commonly arises in problems involving the use of seismic rays to determine acoustic velocity structure.

11.4 TOMOGRAPHY AND CONTINUOUS INVERSE THEORY

The term "tomography" has come to be used in geophysics almost synonymously with the term "inverse theory." Tomography is derived from the Greek word *tomos*, that is, slice, and denotes forming an image of an object from measurements made from slices (or rays) through it. We consider tomography a subset of inverse theory, distinguished by a special form of the data kernel that involves measurements made along rays. The model function in tomography is a function of two or more variables and is related to the data by

$$d_i = \int_{C_i} m[x(s), y(s)] \mathrm{d}s \tag{11.19}$$

Here, the model function is integrated along a curved ray C_i having arc length s. This integral is equivalent to the one in a standard continuous problem

(Equation (11.2)) when the data kernel is $G_i(x,y) = \delta\{x(s) - x_i[y(s)]\}\mathrm{d}s/\mathrm{d}y$, where $\delta(x)$ is the Dirac delta function:

$$d_i = \int\int m(x,y)\delta\{x(s) - x_i[y(s)]\}\frac{\mathrm{d}s}{\mathrm{d}y}\mathrm{d}x\mathrm{d}y = \int_{C_i} m[x(s), y(s)]\mathrm{d}s \qquad (11.20)$$

Here x is supposed to vary with y along the curve C_i and y is supposed to vary with arc length s.

While the tomography problem is a special case of a continuous inverse problem, several factors limit the applicability of the formulas of the previous sections. First, the Dirac delta functions in the data kernel are not square integrable so that the overlap integrals S_{ij} (Equation (11.7)) have nonintegrable singularities at points where rays intersect. Further, in three-dimensional cases, the rays may not intersect at all so that all the S_{ij} may be identically zero. Neither of these problems is insurmountable, and they can be overcome by replacing the rays with tubes of finite cross-sectional width. (Rays are often an idealization of a finite-width process anyway, as in acoustic wave propagation, where they are an infinitesimal wavelength approximation.) Since this approximation is equivalent to some statement about the smoothness of the model function $m(x, y)$, it often suffices to discretize the continuous problem by dividing it into constant m subregions, where the subregions are large enough to guarantee a reasonable number containing more than one ray. The discrete inverse problem is then of the form $d_i = \sum_j G_{ij}m_j$, where the data kernel G_{ij} gives the arc length of the ith ray in the jth subregion. The concepts of resolution and variance, now interpreted in the discrete fashion of Chapter 4, are still applicable and of considerable importance.

11.5 TOMOGRAPHY AND THE RADON TRANSFORM

The simplest tomography problem involves straight-line rays and a two-dimensional model function $m(x, y)$ and is called Radon's problem. By historical convention, the straight-line rays C_i in Equation (11.19) are parameterized by their perpendicular distance u from the origin and the angle θ (Figure 11.2) that the perpendicular makes with the x-axis. Position (x,y) and ray coordinates (u,s), where s is arc length, are related by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} u \\ s \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} u \\ s \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \qquad (11.21)$$

The tomography problem is then

$$d(u,\theta) = \int_{-\infty}^{+\infty} m(x = u\cos\theta - s\sin\theta, y = u\sin\theta + s\cos\theta)\mathrm{d}s \qquad (11.22)$$

In realistic experiments, $d(u, \theta)$ is sampled only at discrete points $d_i = d(u_i, \theta_i)$. Nevertheless, much insight can be gained into the behavior of Equation (11.22) by regarding (u, θ) as continuous variables. Equation (11.22) is then an integral

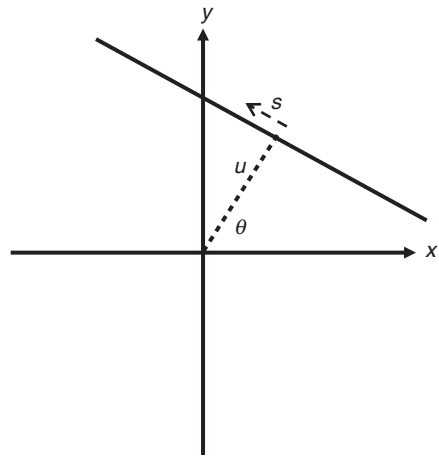

FIGURE 11.2 The Radon transform is performed by integrating a function of (x,y) along straight lines (bold) parameterized by the arc length, s; perpendicular distance, u; and angle, θ.

transform that transforms variables (x,y) into two new variables (u,θ) and is called a *Radon* transform.

11.6 THE FOURIER SLICE THEOREM

The Radon transform is similar to another integral transform, the Fourier transform, which transforms spatial position x into spatial wave number k_x

$$\hat{f}(k_x) = \int_{-\infty}^{+\infty} f(x)\exp(ik_x x)\mathrm{d}x \quad \text{and} \quad f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(k_x)\exp(-ik_x x)\mathrm{d}k_x \tag{11.23}$$

In fact, the two are quite closely related, as can be seen by Fourier transforming Equation (11.22) with respect to $u \rightarrow k_u$:

$$\hat{d}(k_u,\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} m(u\cos\theta - s\sin\theta, u\sin\theta + s\cos\theta)\mathrm{d}s\exp(ik_u u)\mathrm{d}u \tag{11.24}$$

We now transform the double integral from $\mathrm{d}s\ \mathrm{d}u$ to $\mathrm{d}x\ \mathrm{d}y$, using the fact that the Jacobian determinant $|\det[\partial(x, y)/\partial(u, s)]|$ is unity (see Equation (11.21)):

$$\begin{aligned}\hat{d}(k_u,\theta) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} m(x,y)\exp(ik_u\, x\cos\theta + ik_u\, y\sin\theta)\mathrm{d}x\mathrm{d}y \\ &= \hat{\hat{m}}(k_x = k_u\cos\theta,\ k_y = k_u\sin\theta)\end{aligned} \tag{11.25}$$

This result, called the *Fourier slice theorem*, provides a method of inverting the Radon transform. The Fourier-transformed quantity $\hat{d}(k_u,\theta)$ is simply the

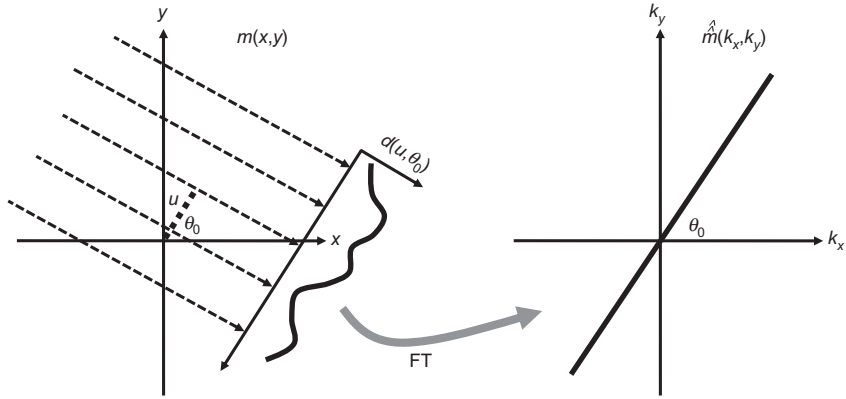

FIGURE 11.3 (Left) The function $m(x,y)$ is integrated along a set of parallel lines (dashed) in a Radon transform to form the function $d(u,\theta_0)$. This function is called the *projection* of $m(x,y)$ at the angle θ_0. (Right) The Fourier slice theorem states that the Fourier transform (FT) of the projection is equal to the Fourier-transformed image evaluated along a line (bold) of angle θ_0 in the (k_x,k_y) plane.

Fourier-transformed image $\hat{\hat{m}}(k_x,\ k_y)$ evaluated along radial lines in the (k_x,k_y) plane (Figure 11.3). If the Radon transform is known for all values of (u,θ), then the Fourier-transformed image is known for all (k_x,k_y), and since the Fourier transform can be inverted uniquely, the image itself is known for all (x,y).

Since the Fourier transform and its inverse are unique, the Radon transform can be uniquely inverted if it is known for all possible (u,θ). Further, the Fourier slice theorem can be used to invert the Radon transform in practice by using discrete Fourier transforms in place of integral Fourier transforms. However, u must be sampled sufficiently evenly that the $u \rightarrow k_u$ transform can be performed and θ must be sampled sufficiently finely that $\hat{\hat{m}}(k_x,\ k_y)$ can be sensibly interpolated onto a rectangular grid of (k_x,k_y) to allow the $k_x \rightarrow x$ and $k_y \rightarrow y$ transforms to be performed (see the discussion in *MatLab* script gda11_01). An example of the use of the Fourier slice theorem to invert tomography data is shown in Figure 11.4.

11.7 CORRESPONDENCE BETWEEN MATRICES AND LINEAR OPERATORS

The continuous function $m(x)$ is the continuous analog to the vector $\mathbf{m}$. What is the continuous analog to $\mathbf{Lm}$, where $\mathbf{L}$ is a matrix? Let us call it $\mathcal{L}m$, that is, $\mathcal{L}$ *operating* on the function $m(x)$. In the discrete case, $\mathbf{Lm}$ is another vector, so by analogy $\mathcal{L}m$ is another function. Just as $\mathbf{L}$ is linear in the sense that $\mathbf{L}(\mathbf{m}^A+\mathbf{m}^B)=\mathbf{Lm}^A+\mathbf{Lm}^B$, we would like to choose $\mathcal{L}$ so that it is linear in the sense that $\mathcal{L}(m^A+m^B)=\mathcal{L}m^A+\mathcal{L}m^B$. A hint to the identity of $\mathcal{L}$ can be drawn from the fact that a matrix can be used to approximate derivatives and integrals. Consider, for example,

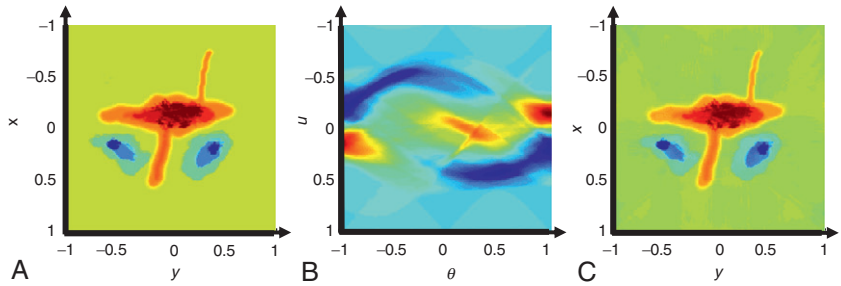

FIGURE 11.4 (A) A test image $m^{\text{true}}(x,y)$ of 256×256 discrete values, or pixels. This synthetic image depicts a hypothetical magma chamber beneath a volcano. (B) The Radon transform $d(u,\theta)$ of the image in (A), also evaluated at 256×256 points. (C) The image $m^{\text{est}}(x,y)$ reconstructed from its Radon transform by direct application of the Fourier slice theorem. Small errors in the reconstruction arise from the interpolation of the Fourier-transformed image onto a rectangular grid. *MatLab* script gda11_02.

$$\mathbf{L}^{\text{A}} = \frac{1}{\Delta x}\begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{L}^{\text{B}} = \Delta x \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \tag{11.26}$$

Here, $\mathbf{L}^{\text{A}}\mathbf{m}$ is the finite difference approximation to the derivative $\mathrm{d}m/\mathrm{d}x$ and $\mathbf{L}^{\text{B}}\mathbf{m}$ is the Reimann sum approximation to the indefinite integral

$$\int_0^x m(x')\mathrm{d}x' \tag{11.27}$$

Thus, $\mathcal{L}$ can be any linear combination of integrals and derivatives (a *linear operator*, for short) (Lanczos, 1961). In the general multidimensional case, where $m(\mathbf{x})$ is a function of a N-dimensional position vector $\mathbf{x}$, $\mathcal{L}$ is built up of partial derivatives and volume integrals. However, for simplicity, we restrict ourselves to the one-dimensional case here.

Now let us consider whether we can define an inverse operator $\mathcal{L}^{-1}$ that is the continuous analog to the matrix inverse $\mathbf{L}^{-1}$. By analogy to $\mathbf{L}^{-1}\mathbf{L}\mathbf{m} = \mathbf{m}$, we want to choose $\mathcal{L}^{-1}$ so that $\mathcal{L}^{-1}\mathcal{L}m(x) = m(x)$. Actually, the derivative matrix $\mathbf{L}^{\text{A}}$ (Equation (11.26)) is problematical in this regard, for it is one row short of being square, and thus has no inverse. This corresponds to a function being determined only up to an integration constant when its derivative is known. This problem can be patched by adding to it a top row

$$\mathbf{L}^{\text{C}} = \frac{1}{\Delta x}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \tag{11.28}$$

that fixes the value of m_1, that is, by specifying the integration constant. Thus, $\mathcal{L}$ is not just a linear operator, but rather a linear operator plus one or more associated *boundary conditions*. It is easy to verify that $\mathbf{L}^{\mathrm{C}}\mathbf{L}^{\mathrm{B}} = \mathbf{I}$. The analogous continuous result that the derivative is the inverse operator of the derivative

$$m(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_0^x m(x')\mathrm{d}x' \tag{11.29}$$

is well known and is called the *fundamental theorem of calculus*. Note that a linear differential equation can be written $\mathcal{L}m(x) = f(x)$. Its solution in terms of the inverse operator is $m(x) = \mathcal{L}^{-1}f(x)$. But its solution can also be written as a Green function integral

$$m(x) = \int_{-\infty}^{+\infty} F(x,\xi)f(\xi)\mathrm{d}\xi = \mathcal{L}^{-1}f(x) \tag{11.30}$$

where $F(x,\xi)$ solves $\mathcal{L}\,F(x,\xi) = \delta(x-\xi)$. Hence, the inverse operator to a differential operator is the Green function integral.

Another important quantity ubiquitous in the formulas of inverse theory is the dot product between two vectors; written here generically as $s = \mathbf{a}^{\mathrm{T}}\mathbf{b} = \sum_i a_i b_i$ where s is a scalar. The continuous analog is the integral

$$s = \int a(\mathbf{x})b(\mathbf{x})\mathrm{d}^N x = (a,b) \tag{11.31}$$

where s is a scalar and $\mathrm{d}^N x$ is the volume element, and the integration is over the whole space of $\mathbf{x}$. This integral is called the *inner product* of the functions $a(x)$ and $b(x)$ and is abbreviated $s = (a,b)$. As before, we restrict the discussion to the one-dimensional case:

$$s = \int_{-\infty}^{+\infty} a(x)b(x)\mathrm{d}x = (a,b) \tag{11.32}$$

Many of the dot products that we encountered earlier in this book contained matrices, for example, $[\mathbf{Aa}]^{\mathrm{T}}\mathbf{b}$ where $\mathbf{A}$ is a matrix. The continuous analog is an inner product containing a linear operator, that is $(\mathcal{L}a,b)$.

An extremely important property of the dot product $[\mathbf{Aa}]^{\mathrm{T}}\mathbf{b}$ is that it can also be written as $\mathbf{a}^{\mathrm{T}}[\mathbf{Bb}]$ with $\mathbf{B} = \mathbf{A}^{\mathrm{T}}$. We can propose the analogous relationship for linear operators:

$$(\mathcal{L}^{\mathrm{A}}a,b) = (a,\mathcal{L}^{\mathrm{B}}b) \tag{11.33}$$

The question then is what is the relationship between the two linear operators $\mathcal{L}^{\mathrm{A}}$ and $\mathcal{L}^{\mathrm{B}}$, or put another way, what is the continuous analog of the transpose of matrix? As before, we sill start off by merely giving the answer a name, the *adjoint*, and a symbol, $\mathcal{L}^{\dagger}$

$$(\mathcal{L}a,b) = (a,\mathcal{L}^{\dagger}b) \tag{11.34}$$

Several approaches are available for determining the $\mathcal{L}^{\dagger}$ corresponding to a particular $\mathcal{L}$. The most straightforward is to start with the definition of the inner

product. For instance, if $\mathcal{L}=c(x)$, where $c(x)$ is an ordinary function, then $\mathcal{L}^\dagger=c(x)$, too, since

$$\int_{-\infty}^{+\infty}(ca)b\mathrm{d}x=\int_{-\infty}^{+\infty}a(cb)\mathrm{d}x \tag{11.35}$$

When $\mathcal{L}=\mathrm{d}/\mathrm{d}x$, we can use integration by parts to find $\mathcal{L}^\dagger$:

$$\int_{-\infty}^{+\infty}\frac{\mathrm{d}a}{\mathrm{d}x}b\mathrm{d}x=ab|_{-\infty}^{+\infty}-\int_{-\infty}^{+\infty}a\frac{\mathrm{d}b}{\mathrm{d}x}\mathrm{d}x \tag{11.36}$$

So $\mathcal{L}^\dagger=-\mathrm{d}/\mathrm{d}x$, as long as the functions approach zero as x approaches $\pm\infty$. This same procedure, applied twice, can be used to show that $\mathcal{L}=\mathrm{d}^2/\mathrm{d}x^2$ is *self-adjoint*, that is, it is its own adjoint. As another example, we derive the adjoint of the indefinite integral $\int_{-\infty}^{x}\mathrm{d}\xi$ by first writing it as

$$\int_{-\infty}^{x}a(\xi)\mathrm{d}\xi=\int_{-\infty}^{+\infty}H(x-\xi)a(\xi)\mathrm{d}\xi \tag{11.37}$$

where $H(x-\xi)$ is the Heaviside step function, which is unity if $x>\xi$ and zero if $x<\xi$. Then

$$\begin{aligned}(\mathcal{L}a,b)&=\int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{x}a(\xi)\mathrm{d}\xi\right\}b(x)\mathrm{d}x\\&=\int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty}H(x-\xi)a(\xi)\mathrm{d}\xi\right\}b(x)\mathrm{d}x\\&=\int_{-\infty}^{+\infty}a(\xi)\left\{\int_{-\infty}^{+\infty}H(x-\xi)b(x)\mathrm{d}x\right\}\mathrm{d}\xi\\&=\int_{-\infty}^{+\infty}a(\xi)\left\{\int_{\xi}^{+\infty}b(x)\mathrm{d}x\right\}\mathrm{d}\xi=(a,\mathcal{L}^\dagger b)\end{aligned} \tag{11.38}$$

Hence, the adjoint of $\int_{-\infty}^{x}\mathrm{d}\xi$ is $\int_{x}^{+\infty}\mathrm{d}\xi$.

Another technique for computing an adjoint is to approximate the operator $\mathcal{L}$ as a matrix, transpose the matrix, and then to "read" the adjoint operator back by examining the matrix. In the cases of the integral and first derivative, the transposes are

$$\mathbf{L}^{\mathrm{AT}}=\Delta x\begin{bmatrix}1&1&1&\cdots&1&1\\0&1&1&1&\cdots&1\\&&\ddots&&&\\0&0&0&\cdots&0&1\end{bmatrix}\quad\text{and}\quad\mathbf{L}^{\mathrm{CT}}=\frac{1}{\Delta x}\begin{bmatrix}1&-1&0&\cdots&0&0\\0&1&-1&0&\cdots&0\\&&\ddots&&&\\0&0&0&\cdots&1&-1\\0&0&0&0&0&1\end{bmatrix} \tag{11.39}$$

$\mathbf{L}^{\mathrm{AT}}$ represents the integral from progressively larger values of x to infinity and is equivalent to $\int_{x}^{+\infty}\mathrm{d}\xi$. All but the last row of the $\mathbf{L}^{\mathrm{CT}}$ represents –d/dx, and the

last row is the boundary condition (which has moved from the top of $\mathbf{L}^{\mathrm{C}}$ to the bottom of $\mathbf{L}^{\mathrm{CT}}$). Hence, these results agree with those previously derived.

Adjoint operators have many of the properties of matrix transposes, including

$$\begin{aligned} &\left(\mathcal{L}^{\dagger}\right)^{\dagger} = \mathcal{L} \quad \text{and} \quad (\mathcal{L}^{-1})^{\dagger} = (\mathcal{L}^{\dagger})^{-1} \\ &(\mathcal{L}^{\mathrm{A}} + \mathcal{L}^{\mathrm{B}})^{\dagger} = (\mathcal{L}^{\mathrm{B}})^{\dagger} + (\mathcal{L}^{\mathrm{A}})^{\dagger} \quad \text{and} \quad (\mathcal{L}^{\mathrm{A}}\mathcal{L}^{\mathrm{B}})^{\dagger} = (\mathcal{L}^{\mathrm{B}})^{\dagger}(\mathcal{L}^{\mathrm{A}})^{\dagger} \end{aligned} \tag{11.40}$$

11.8 THE FRÉCHET DERIVATIVE

Previously, we wrote the relationship between the model $m(x)$ and the data d_i as

$$d_i = \int G_i(x) m(x) \mathrm{d}x = (G_i, m) \tag{11.41}$$

But now, we redefine it in terms of perturbations around some reference model $m^{(0)}(x)$. We define $m(x) = m^{(0)}(x) + \delta m(x)$ where $m^{(0)}(x)$ is a reference function and $\delta m(x)$ is a perturbation. If we write $d_i = d_i^{(0)} + \delta d_i$ where $d_i^{(0)} = (G_i, m^{(0)})$ is the data predicted by the reference model, then

$$\delta d_i = \int G_i(x) \delta m(x) \mathrm{d}x = (G_i, \delta m) \tag{11.42}$$

This equation says that a perturbation $\delta m(x)$ in the model causes a perturbation δd_i in the data. This formulation is especially useful in linearized problems, since then the data kernel can be approximate, that is, giving results valid only when $\delta m(x)$ is small.

Equation (11.42) is reminiscent of the standard derivative formula that we have used in linearized discrete problems:

$$\Delta d_i = \sum_{i=1}^{M} G_{ij}^{(0)} \Delta m_j \quad \text{with} \quad G_{ij}^{(0)} = \left. \frac{\partial d_i}{\partial m_j} \right|_{\mathbf{m}^{(0)}} \tag{11.43}$$

The only difference is that $\Delta\mathbf{m}$ is a vector, whereas $\delta m(x)$ is a continuous function. Thus, we can understand $G_i(x)$ in Equation (11.42) an analog to a derivative. We might use the notation

$$G_i(x) = \left. \frac{\delta d_i}{\delta m} \right|_{\mathbf{m}^{(0)}} \tag{11.44}$$

in which case it is called the *Fréchet derivative* of the datum d_i with respect to the model $m(x)$.

11.9 THE FRÉCHET DERIVATIVE OF ERROR

Fréchet derivatives of quantities other than the data are possible. One that is of particular usefulness is the Fréchet derivative of the error E with respect to the model, where the data $d(x)$ is taken to be a continuous variable.

$$E = (d^{\text{obs}} - d, d^{\text{obs}} - d) \quad \text{and} \quad \delta E = E - E^{(0)} = \left(\left. \frac{\delta E}{\delta m} \right|_{\text{m}^{(0)}}, \delta m \right) \tag{11.45}$$

Here the error is the continuous analog to the L_2 norm discrete error $E = (\mathbf{d}^{\text{obs}} - \mathbf{d})^{\text{T}}(\mathbf{d}^{\text{obs}} - \mathbf{d})$. Suppose now that the data are related to the model via a linear operator, $d = \mathcal{L}m$. We compute the perturbation δE due to the perturbation δm as

$$\begin{aligned} \delta E &= E - E^{(0)} = (d^{\text{obs}} - d, d^{\text{obs}} - d) - (d^{\text{obs}} - d^{(0)}, d^{\text{obs}} - d^{(0)}) \\ &= -2(d, d^{\text{obs}}) + (d, d) + 2(d^{(0)}, d^{\text{obs}}) - (d^{(0)}, d^{(0)}) \\ &= -2(d^{\text{obs}} - d^{(0)}, d - d^{(0)}) + (d - d^{(0)}, d - d^{(0)}) \\ &= -2(d^{\text{obs}} - d^{(0)}, \delta d) + (\delta d, \delta d) \approx -2(d^{\text{obs}} - d^{(0)}, \delta d) \\ &= -2(d^{\text{obs}} - d^{(0)}, \mathcal{L}\delta m) \end{aligned} \tag{11.46}$$

Note that we have ignored a term involving the second-order quantity $(\delta d)^2$. Using the adjoint operator $\mathcal{L}^\dagger$, we find

$$\delta E = (-2\mathcal{L}^\dagger(d^{\text{obs}} - d^{(0)}), \delta m) \tag{11.47}$$

which implies that the Fréchet derivative of the error E is

$$\left. \frac{\delta E}{\delta m} \right|_{\mathbf{m}^{(0)}} = -2\mathcal{L}^\dagger(d^{\text{obs}} - d^{(0)}) \tag{11.48}$$

This result can be useful when solving the inverse problem using a gradient method (see Section 9.8). As an example, consider the problem $d = \mathcal{L}m$ where

$$\mathcal{L}m(x) = a\frac{\text{d}}{\text{d}x}m(x) + b\int_{-\infty}^{x} m(x')\text{d}x' \tag{11.49}$$

where a and b are constants. Using the adjoint relationships derived in Section 11.7, we find

$$\mathcal{L}^\dagger d(x) = -a\frac{\text{d}}{\text{d}x}d(x) + b\int_{x}^{-\infty} d(x')\text{d}x' \tag{11.50}$$

and hence the Fréchet derivative of the error E is

$$\left. \frac{\delta E}{\delta m} \right|_{\mathbf{m}^{(0)}} = 2a\frac{\text{d}}{\text{d}x}[d^{\text{obs}}(x) - d^{(0)}(x)] - 2b\int_{x}^{-\infty} [d^{\text{obs}}(x') - d^{(0)}(x')]\,\text{d}x' \tag{11.51}$$

In order to use this result in a numerical scheme, one must discretize it; say by using voxels of width Δx and amplitude $\mathbf{m}$ for the model and $\mathbf{d}$ for the data, both of length $M = N$. Equation (11.51) then yields the gradient $\partial E/m_i$, which can be used to minimize E via the gradient method (Figure 11.5).

11.10 BACKPROJECTION

The formula for the Fréchet derivative of the error E (Equation (11.48)) is the continuous analog of the discrete gradient $\nabla E = -2\mathbf{G}^{\mathbf{T}}(\mathbf{d}^{\text{obs}} - \mathbf{d}^{\text{pre}}) = -2\mathbf{G}^{\mathbf{T}}(\mathbf{d}^{\text{obs}} - \mathbf{Gm})$. In Section 3.4, the discrete gradient was set to zero, leading

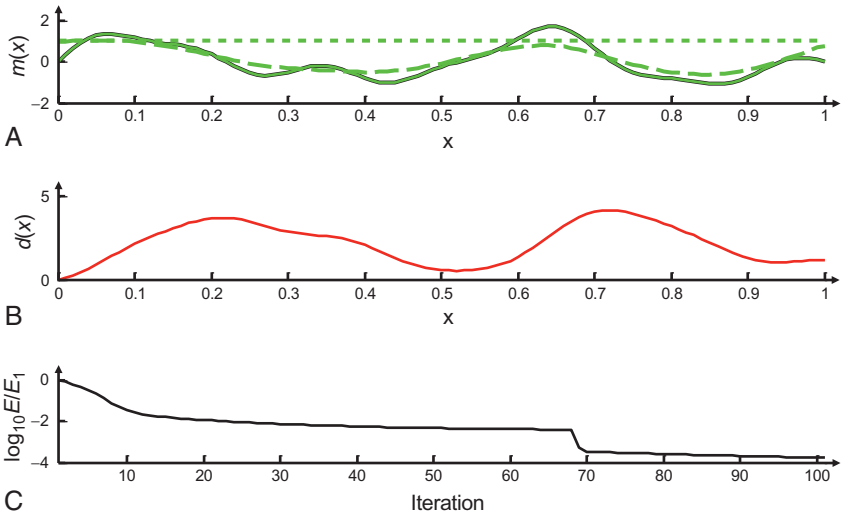

FIGURE 11.5 Example of the solution of a continuous inverse problem using a gradient method to minimize the error E, where an adjoint method is used to compute ∇E. (A) A test function, $m^{\text{true}}(x)$ (green), trial function (dotted green) reconstructed function after 40 iterations (dashed green) and final reconstructed function after 15,890 iterations (green). (B) The data, $d(t)$, satisfies $d(t) = \mathcal{L}m(t)$, where $\mathcal{L}$ is the linear operator discussed in the text. (C) Error E as a function of iteration number, for the first 100 iterations. *MatLab* script gda11_03.

to the least-squares equation for the model parameters, $\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{m} = \mathbf{G}^{\mathrm{T}}\mathbf{d}^{\text{obs}}$. A similar result is achieved in the continuous case:

$$\left.\frac{\delta E}{\delta m}\right|_{\mathbf{m}^{(0)}} = 0 = -2\mathcal{L}^{\dagger}(d^{\text{obs}} - d) = -2\mathcal{L}^{\dagger}(d^{\text{obs}} - \mathcal{L}m) \quad \text{or}$$
$$\mathcal{L}^{\dagger}\mathcal{L}m = \mathcal{L}^{\dagger}d^{\text{obs}} \tag{11.52}$$

Now suppose that $\mathcal{I}$ represents the identity operator, that is, satisfying $m = \mathcal{I}m$ (in one dimension, this operator is $m(x) = \int_{-\infty}^{\infty}\delta(x - x')m(x')\mathrm{d}x'$). Then we can write

$$(\mathcal{L}^{\dagger}\mathcal{L} + \mathcal{I} - \mathcal{I})m = \mathcal{L}^{\dagger}d^{\text{obs}} \quad \text{or}$$
$$m = \mathcal{I}m = \mathcal{L}^{\dagger}d^{\text{obs}} - (\mathcal{L}^{\dagger}\mathcal{L} - \mathcal{I})m \tag{11.53}$$

In the special case that $\mathcal{L}^{\dagger}\mathcal{L} = \mathcal{I}$ (that is, $\mathcal{L}^{\dagger} = \mathcal{L}^{-1}$), the solution is very simple, $m = \mathcal{I}m = \mathcal{L}^{\dagger}d^{\text{obs}}$. However, even in other cases the equation is still useful, since it can be viewed as a recursion relating an old estimate of the model parameters $m^{(i)}$ to a new one, $m^{(i+1)}$

$$m^{(i+1)} = \mathcal{L}^{\dagger}d^{\text{obs}} - (\mathcal{L}^{\dagger}\mathcal{L} - \mathcal{I})m^{(i)} \tag{11.54}$$

If the recursion is started with $m^{(0)} = 0$, the new estimate is

$$m^{(1)} = \mathcal{L}^{\dagger} d^{\text{obs}} \tag{11.55}$$

As an example, suppose that $\mathcal{L}$ is the indefinite integral, so that

$$d^{\text{obs}}(x) = \mathcal{L}\, m(x) = \int_{-\infty}^{x} m(x')\mathrm{d}x' \tag{11.56a}$$

and

$$m^{(1)}(x) = \mathcal{L}^{\dagger}\, d^{\text{obs}}(x) = \int_{x}^{\infty} d^{\text{obs}}(x')\mathrm{d}x' \tag{11.56b}$$

Equation (11.56b) may seem crazy, since an indefinite integral is inverted by taking a derivative, not another integral. Yet this result, while approximate, is nevertheless quite good in some cases, at least up to an overall multiplicative factor (Figure 11.6).

Equation (11.56a) might be considered an ultrasimplified one-dimensional tomography problem, relating acoustic slowness $m(x)$ to traveltime $d^{\text{obs}}(x)$. Note that the ray associated with traveltime $d^{\text{obs}}(x)$ starts at $-\infty$ and ends at x; that is, the traveltime at a point x depends only upon the slowness to the left of x. The formula for $m^{(1)}$ has a simple interpretation: the slowness at x is estimated by summing up all the traveltimes for rays ending at points $x' > x$; that is, summing up traveltimes only for those rays that sample the slowness at x. This process, which carries over to multidimensional tomography, is called *backprojection*. The inclusion, in a model parameter's own estimate, of just those data that are affected by it, is intuitively appealing.

Note that the accuracy of the backprojection depends upon the degree to which $\mathcal{L}^{\dagger}\mathcal{L} - \mathcal{I} = 0$ in Equation (11.54). Some insight into this issue can be

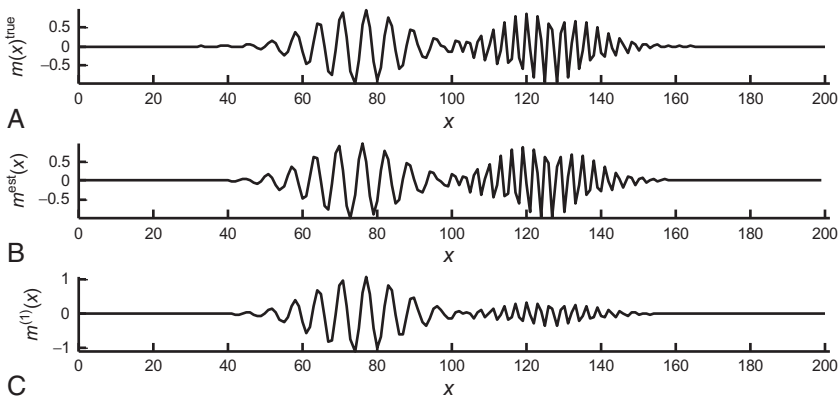

FIGURE 11.6 (A) True one-dimensional model $m^{\text{true}}(x)$. The data satisfy $d^{\text{obs}} = \mathcal{L}m^{\text{true}}$, where $\mathcal{L}$ is the indefinite integral. (B) Estimated model, using $m^{\text{est}} = \mathcal{L}^{-1}\, d^{\text{obs}}$ where $\mathcal{L}^{-1}$ is the first derivative. Note that $m^{\text{est}} = m^{\text{true}}$. (C) Backprojected model $m^{(1)} = \mathcal{L}^{\dagger} d^{\text{obs}}$, where $\mathcal{L}^{\dagger}$ is the adjoint of $\mathcal{L}$. Note that, up to an overall multiplicative factor, $m^{(1)} \approx m^{\text{true}}$. *MatLab* script gda11_04.

gained by examining the singular value decomposition of a discrete data kernel $\mathbf{G} = \mathbf{U}_p \Lambda_p \mathbf{V}_p^T$, which has transpose $\mathbf{G}^T = \mathbf{V}_p \Lambda_p \mathbf{U}_p^T$ and generalized inverse $\mathbf{G}^{-g} = \mathbf{V}_p \Lambda_p^{-1} \mathbf{U}_p^T$. Clearly, the accuracy of the approximation $\mathbf{G}^{-g} = \mathbf{G}^T$ will depend upon the degree to which $\mathbf{\Lambda}_p = \mathbf{\Lambda}_p^{-1}$, that is, to the degree to which the singular values have unit amplitude. The correspondence might be improved by scaling the rows of $\mathbf{G}$ and corresponding elements of $\mathbf{d}^{obs}$ by carefully chosen constants c_i: $G_{ij} \rightarrow c_i G_{ij}$ and $d_i \rightarrow c_i d_i$ so that the singular values are of order unity. This analysis suggests a similar scaling of the continuous problem, $\mathcal{L} \rightarrow c(x)\mathcal{L}$ and $d^{obs}(x) \rightarrow c(x) d^{obs}(x)$, where $c(x)$ is some function. In tomography, a commonly used scaling is $c(x) = 1/L(x)$, where $L(x)$ is the length of ray x. The transformed data d^{obs}/L represents the average slowness along the ray. The backprojection process sums up the average slowness of all rays that interact with the model parameter at point x. This is a bit counterintuitive; one might expect that averaging the averages, as contrasted to summing the averages, would be more appropriate. Remarkably, the use of the summation introduces only long-wavelength errors into the image. An example of two-dimensional back projection is shown in Figure 11.7.

11.11 FRÉCHET DERIVATIVES INVOLVING A DIFFERENTIAL EQUATION

Equation (11.42) links model parameters directly to the data. Some inverse problems are better analyzed when the link is indirect, through a field, $u(x)$. The model parameters are linked to the field, and the field to the data. Consider, for example, a problem in ocean circulation, where the model parameters $m(x)$ represent the force of the wind, the field $u(x)$ represents the velocity of the ocean water and the data d_i represent the volume of water transported from one ocean to another. Wind forcing is linked to water velocity through a fluid-mechanical differential equation, and water transport is related to water velocity through an

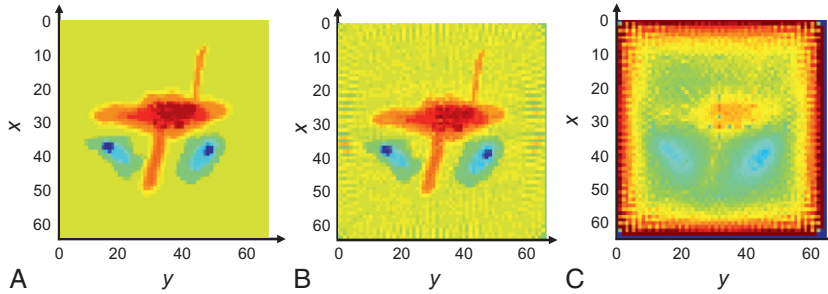

FIGURE 11.7 Example of backprojection. (A) True two-dimensional model, for which travel-times associated with a dense and well-distributed set of rays are measured. (B) Estimated model, using damped least squares. (C) Estimated model, using backprojection. *MatLab* script gda11_05.

integral of the velocity across the ocean-ocean boundary. Such a relationship has the form

$$\mathcal{L}u(x) = m(x) \quad \text{and} \quad \mathcal{L}\delta u(x) = \delta m(x) \tag{11.57}$$

$$d_i = (h_i(x), u(x)) \quad \text{and} \quad \delta d_i = (h_i(x), \delta u(x)) \tag{11.58}$$

Here, Equation (11.57) is a differential equation with known boundary conditions and Equation (11.58) is an inner product involving a known function $h_i(t)$. Symbolically, we can write the solution of the differential equation as

$$\begin{aligned} u(x) &= \int F(x,\xi) m(\xi) \mathrm{d}\xi = \mathcal{L}^{-1} m(x) \\ \delta u(x) &= \int F(x,\xi) \delta m(\xi) \mathrm{d}\xi = \mathcal{L}^{-1} \delta m(x) \end{aligned} \tag{11.59}$$

where $F(x,\xi)$ is the Green function. Note that $\mathcal{L}^{-1}$ is a linear integral operator, whereas $\mathcal{L}$ is a linear differential operator. We now combine Equations. (11.57) and (11.58):

$$\delta d_i = (h_i, \delta u) = (h_i, \mathcal{L}^{-1}\delta m) = ((\mathcal{L}^{-1})^{\dagger} h_i, \delta m) = ((\mathcal{L}^{\dagger})^{-1} h_i, \delta m) \tag{11.60}$$

Comparing to Equation (11.42), we find

$$G_i(x) = (\mathcal{L}^{\dagger})^{-1} h_i(x) \quad \text{or} \quad \mathcal{L}^{\dagger} G_i(x) = h_i(x) \tag{11.61}$$

Thus, the scalar field satisfies the equation $\mathcal{L}u(x) = m(x)$ and the data kernel satisfies the adjunct equation $\mathcal{L}^{\dagger} G_i(x) = h_i(x)$. In most applications, the scalar field $u(t)$ must be computed by numerical solution of its differential equation. Thus, the computational machinery is typically already in place to solve the adjoint differential equation and thus to construct the data kernel.

An important special case is when the data is the field $u(x)$ itself, that is, $d_i = u(x_i)$. This choice implies that the weighting function in Equation (11.58) is a Dirac delta function, $h_i(x) = \delta(x - x_i)$ and that the data kernel is the Green function, say $Q(x,x_i)$, to the adjoint equation. Then

$$d_i = (Q(x, x_i), m(x)) \quad \text{with} \quad \mathcal{L}^{\dagger} Q(x, x_i) = \delta(x - x_i) \tag{11.62}$$

As an example, we consider an object that is being heated by a flame. We will use time t instead of position x in this example; the object is presumed to be of a spatially uniform temperature $u(t)$ that varies with time, t. Suppose that temperature obeys the simple Newtonian heat flow equation

$$\mathcal{L}u(t) = \left\{ \frac{\mathrm{d}}{\mathrm{d}t} + c \right\} u(t) = m(t) \tag{11.63}$$

where the function $m(t)$ represents the heating and c is a thermal constant. This equation implies that in the absence of heating, the temperature of the object decays away toward zero at a rate determined by c. We assume that the heating is restricted to some finite time interval, so that temperature satisfies the boundary condition $u(t \to -\infty) = 0$.

The solution to the heat flow equation can be constructed from its Green function $F(t,\tau)$, which represents the temperature at time t due to an impulse of heat at time τ. It solves the equation

$$\left\{\frac{\mathrm{d}}{\mathrm{d}t} + c\right\} F(t,\tau) = \delta(t-\tau) \tag{11.64}$$

The solution to this equation can be shown to be

$$F(t,\tau) = H(t-\tau)\exp\{-c(t-\tau)\} \tag{11.65}$$

Here the function $H(t-\tau)$ is the Heaviside step function, which is zero when $t<\tau$ and unity when $t>\tau$. This result can be verified by first noting that it satisfies the boundary condition and then by checking, through direct differentiation, that it solves the differential equation. The temperature of the object is zero before the heat pulse is applied, jumps to a maximum at the moment of application, and then exponentially decays back to zero at a rate determined by the constant, c (Figure 11.8A). A hypothetical heating function $m(x)$ and resulting temperature $u(t)$ are shown in Figure 11.8B and C, respectively.

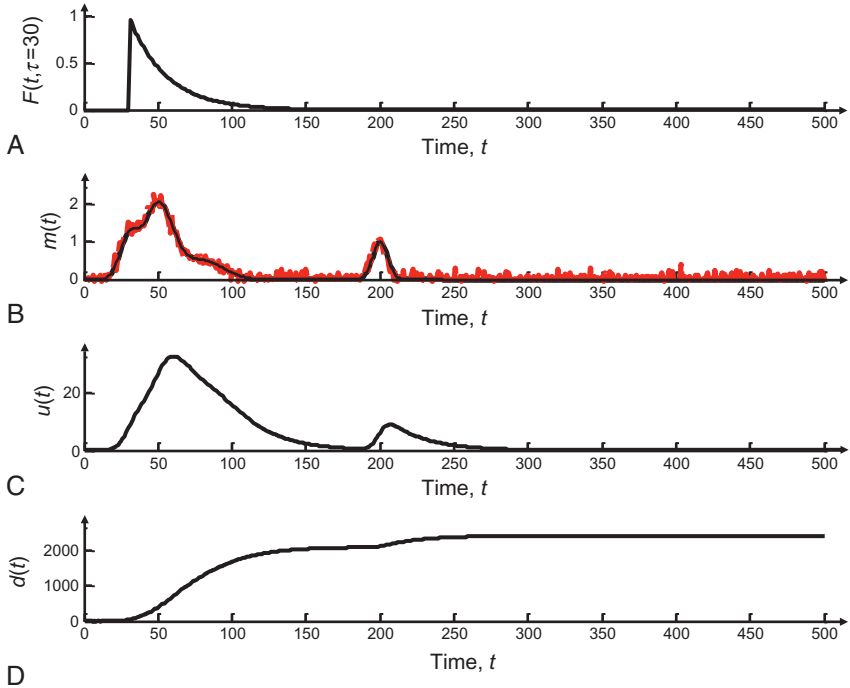

FIGURE 11.8 Example of the solution of a continuous inverse problem involving a differential equation. (A) Green function $F(t,\tau)$ for $\tau=30$. (B) True (black) and estimated (red) heat production function $m(t)$. (C) Temperature $u(t)$, which solves $\mathcal{L}u=m$. (D) Observed data $d(t)$, which is proportional to the integral of $u(t)$. *MatLab* script gda11_06.

We now need to couple the temperature function $u(t)$ to observations, in order to build a complete forward problem. Suppose that a chemical reaction occurs within the object, at a rate that is proportional to temperature. We observe the amount of chemical product $d_i = P(t_i)$, which is given by the integral

$$d_i = P(t_i) = b\int_0^{t_i} u(t)\mathrm{d}t = b\int_0^{\infty} H(t_i - t)u(t)\,\mathrm{d}t = (bH(t_i - t), u(t)) \quad (11.66)$$

where b is the proportionality factor. Thus, $h_i(t) = bH(t_i - t)$, where H is the Heaviside step function. The adjoint equation is

$$\mathcal{L}^{\dagger}u(t) = \left\{-\frac{\mathrm{d}}{\mathrm{d}t} + c\right\}g_i(t) = h_i(t) \quad (11.67)$$

since, as noted in Section 11.7, the adjoint of $\mathrm{d}/\mathrm{d}t$ is $-\mathrm{d}/\mathrm{d}t$ and the adjunct of a constant c is itself. The Green function of the adjoint equation can be shown to be

$$Q(t,\tau) = H(\tau - t)\exp\{+c(t-\tau)\} \quad (11.68)$$

Note that $Q(t,\tau)$ is just a *time-reversed* version of $F(t,\tau)$, a relationship that arises because the two differential equations differ only by the sign of t. The data kernel $G_i(t)$ is given by the Green function integral

$$\begin{aligned} G_i(t) &= \int_0^{\infty} Q(t,\tau)h_i(\tau)\mathrm{d}\tau \\ &= \int_0^{\infty} H(\tau - t)\exp\{c(t-\tau)\}bH(t_i - \tau)\mathrm{d}\tau \\ &= b\int_0^{\infty} H(\tau - t)H(t_i - \tau)\exp\{c(t-\tau)\}\mathrm{d}\tau \\ &= b\int_t^{t_i} \exp\{-c(\tau - t)\}\mathrm{d}\tau \end{aligned} \quad (11.69)$$

The integral is nonzero only in the case where $t_i > t$:

$$G_i(t) = \begin{cases} 0 & t_i \le t \\ -\dfrac{b}{c}\left[\exp\{-c(t_i - t)\} - 1\right] & t_i > t \end{cases} \quad (11.70)$$

The data kernel $G_i(t)$ (Figure 11.9) quantifies the effect of heat applied at time t on an observation made at time t_i. It being zero for times $t > t_i$ is a manifestation of causality; heat applied in the future of a given observation cannot affect it.

The problem that we have been discussing is completely linear; neither the differential equation relating $m(t)$ to $u(t)$ nor the equation relating $u(t)$ to d_i contains any approximations. The data kernel is therefore accurate for perturbations

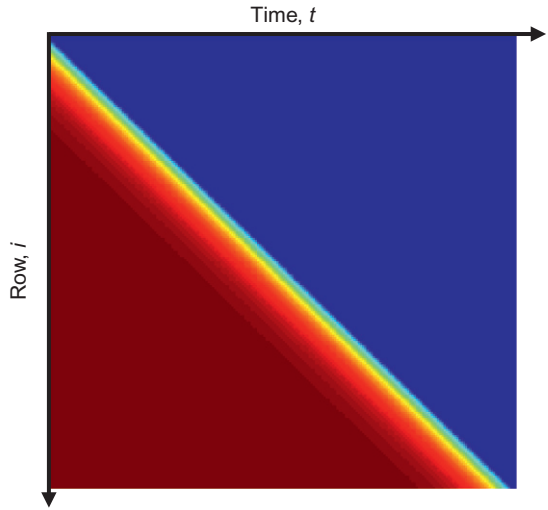

FIGURE 11.9 Data kernel $G_i(t)$ for the continuous inverse problem involving a differential equation. See text for further discussion. *MatLab* script gda11_06.

of any size and hence holds for $m(t)$ and d_i, as well as for the perturbations $\delta m(t)$ and δd_i:

$$d_i = (G_i(t), m(t)) \tag{11.71}$$

We solve the inverse problem—determining the heating through observations of the chemical product—by first discretizing all quantities with a time increment Δt

$$m(t) \rightarrow m_i = m(i\Delta t) \quad \text{and} \quad G_i(t) \rightarrow G_{ij} = \Delta t G_i(j\Delta t) \tag{11.72}$$

so that

$$d_i = (G_i(t), m(t)) \rightarrow \mathbf{d} = \mathbf{Gm} \tag{11.73}$$

Note that a factor of Δt has been included in the definition of $\mathbf{G}$ to account for the one that arises when the inner product is approximated by its Reimann sum. We assume that data is measured at all times, and then solve Equation (11.73) by damped least squares (Figure 11.8B). Some noise amplification occurs, owing to the very smooth data kernel. Rapid fluctuations in heating $m(t)$ have very little effect on the temperature $u(t)$, owing to the diffusive nature of heat, and thus are poorly constrained by the data. Adding smoothness constraints would lead to a better solution.

In the example above, we were concerned with the effect of a perturbation in *forcing* (meaning the function on the right-hand side of the differential equation) on the data. Another common problem is to understand the effect of a

perturbation of a *parameter* in the differential operator, $\mathcal{L}$, on the data. The parameter $c(t)$ in the differential equation

$$\left\{\frac{\mathrm{d}}{\mathrm{d}t}+c(t)\right\}u(t)=f(t) \tag{11.74}$$

would be an example, where the forcing $f(t)$ is now considered a known function. We now write $c(t)=c^{(0)}+\delta c(t)$ and $u(t)=u^{(0)}(t)+\delta u(t)$, where $u^{(0)}(t)$ solves the unperturbed equation

$$\left\{\frac{\mathrm{d}}{\mathrm{d}t}+c^{(0)}\right\}u^{(0)}(t)=f(t) \tag{11.75}$$

The perturbed equation then becomes

$$\begin{aligned}&\left\{\frac{\mathrm{d}}{\mathrm{d}t}+c^{0}+\delta c(t)\right\}\{u^{(0)}(t)+\delta u(t)\}=f(t)\\&\left\{\frac{\mathrm{d}}{\mathrm{d}t}+c^{0}\right\}u^{(0)}(t)+\left\{\frac{\mathrm{d}}{\mathrm{d}t}+c^{0}\right\}\delta u(t)+\delta c(t)u^{(0)}(t)+\delta c(t)\delta u(t)=f(t)\end{aligned} \tag{11.76}$$

After subtracting out the unperturbed equation, ignoring the second-order term and rearranging, we find

$$\left\{\frac{\mathrm{d}}{\mathrm{d}t}+c^{0}\right\}\delta u(t)=-\delta c(t)u^{(0)}(t) \tag{11.77}$$

Thus, the perturbation δc in the parameter c is equivalent to a perturbation in an *unknown* forcing:

$$\delta m(t)=-\delta c(t)u^{(0)}(t) \tag{11.78}$$

We can now use the formalism that we developed above to determine the resulting perturbation in the data.

Adjoint techniques have had extensive application in seismology (Dahlen et al., 2000; Tromp et al., 2005) and in atmospheric and ocean science, where they are associated with the term *data assimilation* (Hall and Cacuci, 1983; Moore et al., 2004).

11.12 PROBLEMS

11.1. This inverse problem is due to Robert Parker. Consider a radially stratified sphere of radius $R=1$ with unknown density $\rho(r)$. Its mass $M=4\pi\int\rho(r)r^2\mathrm{d}r$ and its moment of inertia $I=(\frac{8\pi}{3})\int\rho(r)r^4\mathrm{d}r$ are measured. What is the best averaging kernel that can be designed that is

centered about ½R? Since all the integrals involve only rational functions and all the matrices are 2×2, the problem can be solved analytically. Do as much of the problem as you can, analytically; do the parts that you cannot, numerically.

11.2. Suppose that a function $m(x)$ is discretized in each of two ways: (A) by dividing it into M rectangles, each of height m_i and width Δx; and (B) by representing it as a sum of M overlapping Gaussians, spaced at regular intervals Δx, and with coefficients

$$m(x) = \sum_{i=1}^{M} m_i' \frac{\Delta x}{(2\pi)\sigma} \exp\left\{ -\frac{(x - i\Delta x)^2}{2\sigma^2} \right\}$$

Here, σ^2 is a prescribed variance. Discuss the advantages and disadvantages of each representation. Include the respective effect of damping in the two cases.

11.3. How sensitive are the results of tomography to noise in the data (u, θ)? Run experiments with the Fourier slice script. Try two different types of noise: (A) uncorrelated random noise drawn from a Gaussian distribution and (B) a few large outliers. Comment upon the results.

11.4. Modify the example in Section 11.9 to use a truncated Fourier sine series representation of the model:

$$m^{(0)}(x) = \sum_{n=1}^{K} m_n^{(0)} \sin\left(\frac{n\pi x}{L}\right) \quad \text{and} \quad \delta m(x) = \sum_{n=1}^{K} \delta m_n^{(0)} \sin\left(\frac{n\pi x}{L}\right)$$

where m_n are scalar coefficients, $K = 15$ and $0 < x < L$. You will need to use the chain rule

$$\left.\frac{\delta E}{\delta m_n}\right|_{\mathbf{m}^{(0)}} = \left(\left.\frac{\delta E}{\delta m}\right|_{\mathbf{m}^{(0)}}, \left.\frac{\delta m}{\delta m_n}\right|_{\mathbf{m}^{(0)}} \right)$$

11.5. Suppose that the Green function $F(t,\tau)$ of a particular differential equation $\mathcal{L}^F F(t,\tau) = \delta(t-\tau)$ depends only upon time differences; that is, $F(t,\tau) = F(t-\tau)$. Denote the Green function integral as the linear operator $(\mathcal{L}^F)^{-1}$. (A) Use the formula for the inner product to derive $(\mathcal{L}^F)^{-1\,\dagger}$. Call the function within this operator F'. (B) Suppose the differential equation $\mathcal{L}^{Q\dagger} Q(t,\tau) = \delta\,(t-\tau)$ has Green function $Q(t,\tau)$. Denote the Green function integral as the linear operator $(\mathcal{L}^{Q\dagger})^{-1}$. By comparing $(\mathcal{L}^F)^{-1\,\dagger}$ and $(\mathcal{L}^{Q\dagger})^{-1}$, establish the relationship between F' and Q. Explain your results in words.

REFERENCES

Backus, G.E., Gilbert, J.F., 1968. The resolving power of gross earth data, Geophys. J. Roy. Astron. Soc. 16, 169–205.

Dahlen, F.A., Hung, S.-H., Nolet, G., 2000. Fréchet kernels for finite frequency traveltimes—I. Theory, Geophys. J. Int. 141, 157–174.

Hall, M.C.G., Cacuci, D.G., 1983. Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. J. Atmos. Sci. 40, 2537–2546.

Lanczos, C., 1961. Linear Differential Operators. Van Nostrand-Reinhold, Princeton, NJ.

Moore, A.M., Arango, H.G., Di Lorenzo, E., Cornuelle, B.D., Miller, A.J., Neilson, D.J., 2004. A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint of a regional ocean model. Ocean Modelling 7, 227–258.

Tromp, J., Tape, C., Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, Geophys. J. Int. 160, 195–216.