Geophys. J. Int. (1999) **137,** 175–196

Analysis of potential field data in the wavelet domain

P. Hornby, F. Boschetti and F. G. Horowitz

Australian Geodynamics Cooperative Research Centre and CSIRO Exploration and Mining, PO Box 437, *Nedlands, WA, Australia,* 6009

Accepted 1998 November 16. Received 1998 November 5; in original form 1997 September 12

SUMMARY

Various Green's functions occurring in Poisson potential field theory can be used to construct non-orthogonal, non-compact, continuous wavelets. Such a construction leads to relations between the horizontal derivatives of geophysical field measurements at all heights, and the wavelet transform of the zero height field. The resulting theory lends itself to a number of applications in the processing of potential field data. Some simple, synthetic examples in two dimensions illustrate one inversion approach based upon the maxima of the wavelet transform (multiscale edges). These examples are presented to illustrate, by way of explicit demonstration, the information content of the multiscale edges. We do not suggest that the methods used in these examples be taken literally as a practical algorithm or inversion technique. Rather, we feel that the real thrust of the method is towards physically based, spatially local filtering of geophysical data images using Green's function wavelets, or compact approximations thereto. To illustrate our first steps in this direction, we present some preliminary results of a 3-D analysis of an aeromagnetic survey.

Key words: gravity anomalies, magnetic anomalies, multiscale edge analysis.

1 INTRODUCTION

The collection and analysis of aeromagnetic data represents one of the cheapest forms of large-scale geophysical exploration. It has the particular advantage of allowing relatively easy exploration of remote and hardly accessible areas. As a result, extensive research has been performed in the last decades in order to develop tools that allow the extraction of useful information from potential field data.

Traditionally, the techniques used in potential field analysis could be broadly divided into two classes. In the first class, potential field maps are analysed visually by geoscientists (see for example Telford *et al.* 1976). Image processing tools are used to present the data in a form that facilitates this visual inspection. First and second derivatives, different sun-angle illumination and careful map colouring are typical processing tools that allow geoscientists to discriminate better the features present in the data. The aim of this kind of analysis is usually identification of geological features, such as faults, folds, geological contacts and, more rarely, broad information about the depth and extent of the main geological bodies. Often, the main goal is to produce an approximate 2-D geological map. This approach has the flavour of an art and requires extensive experience. As such, it is partly subjective.

The second kind of approach can be defined generically as inversion. A useful review on some of the methods available for the inversion of potential field data can be found in Blakely (1995). The goal of inversion is typically to ascertain the horizontal and vertical position of main geological bodies. This approach has to face the fundamental limitation of non-uniqueness inherent in the analysis of potential field data (Al-Chalabi 1971; Boschetti *et al.* 1998). In the absence of specific, objective *a priori* information, the selection of one, or some, among the infinite set of possible solutions, again requires subjective judgement.

The two methods differ not only in the goals and in the tools used, but also in the scale of the analysis. Visual inspection and image processing techniques allow the analysis of very large maps with a resolution comparable to the sample spacing, while 'pure' inversion techniques, due to the complexity of the search space and the limitation in computation capabilities, are normally constrained to coarse voxels, or to small segments of a map.

In more recent times, a class of methods intermediate between the visual inspection and 'pure' inversion have gained in popularity. These methods come under the heading of signal processing techniques, and broadly seek approximate inversions through the use of signal processing. In more than one sense, these methods are analogous to migration algorithms in seismic processing. First, they seek to process the signal in order to make a transformation of the signal 'look like' a geological picture of the main features. Second, the signal processing is motivated by the physics of the signal propagation through space. An example is 'depth slicing' or separation filtering, in its various forms (for example Cordell 1985; Jacobsen 1987). The wavelet methods outlined below fit into this middle-of-the-road approach and, indeed, are very similar in their use of upward continuation (field propagation by convolution with the Green's function). However, this similarity should not be allowed to occlude a fundamental difference in our approach. In the present work, we are assuming that the source distribution is deterministic, and we seek specific features in this source structure. This is very different from assuming a statistical ensemble of sources, and leads to different interpretations of the same filters. In particular, the wavelet theory makes precise statements about the decay rates of features in the 'depth-slice' domain.

Our description of multiscale edge detection by wavelet analysis follows in almost every respect the development outlined by Mallat & Zhong (1992). We will show that wavelet analysis is well suited to potential field theory and that fundamental equations, particularly upward continuation, have a very elegant and compact form in the wavelet domain. The upshot is the method's ability to locate the edges of geological bodies, allowing automatic production of the type of map commonly called a 'worm diagram' or skeletonization. Wavelet analysis not only offers fast algorithms for detecting edges in images, but also, by analysis of the wavelet coefficients at different scales, allows the characterization of the structures present in the signal's source. Indeed, an appropriate choice of wavelet can lead to *quantitative* information about the depth and characteristics of causative bodies, and consequently can be used in an inversion procedure. This has also recently been pointed out by Moreau *et al.* (1997), in whose work precisely the same wavelet transforms as we shall describe have been used to locate point sources, and to estimate the degree of homogeneity (structural index) of point sources.

We present 2-D examples using synthetic data that demonstrate, additionally, the recovery of the dimensions and even density contrast of a source anomaly, using a single multiscale edge. Of course the conditions of the simulation are ideal; however, this is good supporting evidence for our more modest claim that multiscale edges will yield practical feature-based filtering techniques, and allow recovery of limited information on source type and depth in real data.

Edge detection applied to potential field data has been explored by other authors (for example Blakely & Simpson 1986). The strength of this new approach is the use of a multiscale edge analysis and the possibility that it offers in the framework of an inversion scheme. The fundamental theory linking the multiscale edges, wavelet transforms and the Green's function of the physical theory is far-reaching, as its applications include full 3-D inversion, data compression, de-noising, enhancement/elimination of specific features and interpolation, all within a single, rigorous framework. The analysis of these applications is ongoing, and we hope to report more of these matters at a later date. Our goal in this paper is an elementary introduction to the theory for the wavelet analysis of potential field data, which relies more upon demonstration of its application to simple 2-D and 3-D examples than upon *rigor mathematicus.*

2 MULTISCALE WAVELET ANALYSIS

2.1 Basic concepts

In a signal or image, it is the points of sharp variation and irregular features that often carry most of the information. In images, for example, spatially coherent discontinuities in the pixels' intensity allow the visual recognition of objects and main patterns. Accordingly, the automatic detection of such features has been the object of intensive research. In the image processing community, the concept of multiscale edge detection has long been used. In 1-D signals, 'edges' are characterized as *points* of extreme value in the first derivative. If one transforms the signal in some way to eliminate or damp features smaller than a given scale s, the resulting signal is often referred to as the signal at scale s. One can seek derivative extrema (edges) in all such scaled signals. These extrema, and their associated magnitudes (as functions of scale), are generally termed *multiscale edges*, and, in the case of 1-D signals, often form a 1-D subset of the space-scale (x, s) plane (see, for example, Mallat 1991). As we shall see, such edges correspond to features in the wavelet transform of the signal.

Sometimes, stationary points of the first derivative corresponding to zero crossings of the second derivative define what is meant by multiscale edges. Such edges also go by the name of 'fingerprints', a guise in which they have already appeared in the geological literature (for example Piech & Piech 1990). However, Mallat & Zong (1992) point out that simple inflection points can be a misleading characterization of edges, and that the subset of second-derivative zero crossings corresponding to the maxima of the absolute value of the first derivative might be a more pleasing definition of edge from a perceptual point of view. In the following, we shall not be particularly pedantic about precisely which definition of multiscale edge we refer to in our general discussions, although we usually follow the convention of Mallat & Zong when presenting results.

In two dimensions, there is a greater choice for the characterization of 'edge'. For example, the maximum magnitude of the gradient along gradient streamlines (paths with gradient as tangent), is one analogue of the 1-D maximum of absolute value of gradient. Similarly, the zeros of the (2-D) Laplacian are analogous to the 1-D inflection points. In the following we shall follow Mallat & Zong, and use the maximum gradient along streamline characterization. This seems more natural from the perceptual point of view; however, it is somewhat at odds with mathematical convenience, which militates for an edge definition expressed in terms of the zero crossings (i.e. level crossings) of some linear differential operator.

Since the main features in an image are represented by its edges, an open issue in image processing research in the last decade has been the image reconstruction question: 'Is the information contained in the edges sufficient to reconstruct the entire original image?' (Hummel & Moniot 1989; Yuille & Poggio 1986). Also note that, in our application domain, we are asking for the reconstruction of the *image* (the original data) from the multiscale edges, not the reconstruction of the *source*. The latter is known to be impossible. In addition, we wish to do this image reconstruction from *just* the edges, and the values of a wavelet transform on those edges, *not* from the whole wavelet transform. The latter inverse transformation is presented in Appendix B.

The answer to the reconstruction question for completely general images and commonly used Gaussian wavelets is 'no'. However, Mallat & Zhong (1992) showed that a very good approximation of the original image can be reconstructed simply from the information contained in the location and amplitude of the local maxima of the wavelet coefficients, that is, in the position and 'sharpness' of the edges at different scales. However, in the case where a well-defined physical process (a potential field measurement) generates an image, a stronger result may obtain, especially if precisely the correct wavelets are used for the definition of the edges.

In the case we consider, the wavelet transform is (apart from a stretching of the scale-coordinate) a harmonic function of the space-scale half-space. Thus the values of the transform on the entire space-scale half-space can be constructed from its values on a suitable surface (a Dirichlet problem). It then remains to show that the multiscale edges form a suitable 'boundary' for such a problem, a consideration which at present we feel is related to the infinite differentiability of, and maximum principle for, harmonic functions (Hummel & Moniot 1989; Mikhailov 1978).

Just as important is the need for a 'boundary' surface that leads to a *stable* solution. For example, knowledge of the values of a harmonic function on a horizontal plane is sufficient to determine the values everywhere, but downward continuation is nevertheless unstable, so the theoretical proof of the existence of a solution can at times be useless from a practical point of view. The definition used to define an edge is important also in this respect, since the resulting multiscale edges must not only furnish a suitable boundary for the solution of the Dirichlet problem, but must also be sufficiently redundant in their information content to furnish a stable inversion for the original image. Ironically, it is this instability of the solution from a single boundary, together with the possibility of using many boundaries (many multiscale edges) as 'regularizing constraints' on the solution process, that makes the signal processing applications of the multiscale edges plausible. The accessibility of results pertaining to the topology of the multiscale edges, and this stability/redundancy issue, makes the perceptually less appealing 'zero of Laplacian' definition of edge mathematically appealing.

Nevertheless, should this reconstruction conjecture prove not to be true, the approximate reconstruction property appears to be good enough to allow image compression, de-noising, feature isolation and interpolation. Preliminary investigations into the reconstruction wavelets for continuous space and discrete scale are promising (Appendix B), and we hope to report more fully on these matters in due course. Here, we will concentrate upon the elementary application to potential field data, and show how the elements of the theory relate to interpretation and inversion.

2.2 Multiscale edge detection

Let us briefly review the main concepts behind multiscale edge detection for 1-D signals. The application of the analysis to potential field data in two and three dimensions will be described in the following sections.

Let $\theta(x)$ be a non-negative differentiable function that converges to zero at infinity sufficiently quickly that the following manipulations are valid. Suppose, in addition, that its integral is unity, then the same conditions hold true of the scaled functions

$$\theta_s(x) = (1/s)\theta(x/s)\,, \qquad s > 0\,. \tag{2.1}$$

Moreover, define

$$\psi(x) = \theta_{,1}(x) = D_x\theta(x)\,, \tag{2.2}$$

whose integral is consequently zero. (Note that D_x denotes differentiation with respect to x, and that, in preparation for more than one dimension, we have already introduced the comma subscript convention to denote differentiation in the classical tensorial style.) The functions θ and ψ can be used as smoothing function and corresponding wavelet respectively. An often-used example pair is the Gaussian and its derivative.

The wavelet transform of a function $f(x)$ is defined by

$$W[f](s, x) = [f{*}\psi_s](x)\,, \tag{2.3}$$

where $\psi_s(x)$ is a dilated version of the wavelet defined by

$$\psi_s(x) = (1/s)\psi(x/s)\,, \tag{2.4}$$

$*$ denotes convolution, and s represents the scale. From (2.1) it follows that

$$W[f](s, x) = [f{*}(sD_x\theta_s)](x) = sD_x[f{*}\theta_s](x)\,. \tag{2.5}$$

Eq. (2.5) shows that, except for a multiplicative factor of s, the wavelet transform $W[f]$ is the first derivative of $[f{*}\theta_s]$. Thus the maxima of $|W[f](s, x)|$ (in x) correspond to rapid variations in $[f{*}\theta_s]$ and hence can be interpreted as regions of rapidly changing intensity; that is, edges in the signal $[f{*}\theta_s]$. Now, $[f{*}\theta_s]$ is just the original signal f blurred or averaged by the filter kernel θ_s. However θ_s averages f over its values within a span of length (approximately) s. Consequently, the local maxima of $|W[f]|$ correspond to edges in the signal after blurring at scale s. These edges, which become a function of scale s, are termed *multiscale edges*. This gives us a mathematical characterization of the notion of an edge in terms of a feature of the wavelet transform of the function, and thereby allows a quite thorough study of what was once an empirical construction used in the analysis of images.

2.3 Lipschitz regularity and wavelet transforms

A feature of multiscale edge detection in the wavelet domain, crucial for the following discussion, is the possibility of characterizing discontinuities and sharp structures in the data by studying the behaviour of the local extrema of the wavelet transform at different scales. In particular, the scale dependence of the wavelet transform depends on the Lipschitz *exponent* of the original signal.

Let us review the concept of the Lipschitz exponent, and its geophysical significance. We will concentrate upon 'well-behaved' functions defined on sets of integer dimension, and refer to Appendix A for a more complete discussion of less regular situations. We recall that, given an $\alpha \in (0, 1]$, a function is said to be uniform Lipschitz α in the interval (a, b) if there exists a constant C (independent of x_i) such that, for all $x_0 \in (a, b)$ and $x_1 \in (a, b)$,

$$|f(x_0) - f(x_1)| \leq C|x_0 - x_1|^{\alpha}\,. \tag{2.6a}$$

Lipschitz uniform regularity of $f(x)$ [denoted as LUR(f)] is defined as the maximum α such that $f(x)$ is uniform Lipschitz α. We do mean the maximum here, not the supremum, and the maximum may not exist. This is slightly at odds with the usual definition. However, we wish to treat the case when the maximum does not exist as a special case, since such a situation does arise in the following, and we prefer to modify (2.6a) to deal with this important case.

In Appendix A we show that f can be made continuous at the end-points a and b, and give additional conditions under which f can be differentiated to yield a conventional function. If f can be differentiated, the derivative $D_x f$ can diverge at a finite number of points, but the singularities are integrable, and the exponent of the divergence is bounded by $\alpha - 1$. In the case mentioned above, when the maximum exponent does not exist, the usual cause is integrable logarithmic singularities in $D_x f$.

If $\alpha = 0$ then (2.6a) confines the values of f to an interval of length C, but f need not be continuous. For example, a step function satisfies (2.6a) with exponent zero. In the following discussion, we constrain the number and type of jump singularities, and allow the derivative to acquire a finite number of Dirac delta function contributions to model these jumps. More complex situations are considered in Appendix A.

For our present application, it is possible to give meaning to negative uniform Lipschitz regularity as follows. If f does not satisfy (2.6a) for any $\alpha \in [0, 1]$, but f is nevertheless integrable, then we define the function

$$F(x) = \int_{[a,x]} f(t)\, dt\,.$$

Suppose that $\mathrm{LUR}(F) \in [0, 1)$, then we define $\mathrm{LUR}(f) = \mathrm{LUR}(F) - 1$. Note that such an f generally does not satisfy (2.6) for the negative value of α. However, if f is a function in the conventional sense, then $\mathrm{LUR}(f)$ does bound the exponent of any divergence in f. The case $\mathrm{LUR}(f) = -1$ is intended to assert that f is a density with Dirac delta contributions. The definition of negative Lipschitz regularity in terms of the primitive function F is appropriate for our present application because f will be convolved with harmonic functions. Our interest is in the Lipschitz regularity of such convolutions and its relationship to source structure.

The above discussion was for functions and measures of a single variable. If the x_i in (2.6a) become n-dimensional vectors, $\mathbf{x}_i$, then the modulus on the right of (2.6a) is replaced by the Euclidean norm of the difference (for example eq. A6). If $\alpha \in (0, 1]$ then f again extends to a continuous function on the closure of the domain and the previous discussion now applies to each line segment in the domain. If $\mathrm{LUR}(f) = 0$ in n dimensions, we will only consider the case when the jumps in f occur along some regular, rectifiable $(n-1)$-dimensional subset.

When f ceases to satisfy the Lipschitz condition (2.6a) in n dimensions, but is still an integrable density, then we can once again define what we mean by $\mathrm{LUR}(f)$ in a way that is consistent with our purpose. Let us defer the discussion of this matter entirely to Appendix A, and focus upon the geological interpretation that arises. The correspondence to be presented is somewhat simpler than is actually the case, but it is a good guide to the significance of the following, and a fairly neat summary of the rather technical description in Appendix A.

The geophysical interpretation of the Lipschitz regularity is (roughly) as follows. Positive Lipschitz exponents correspond to smooth variations. A zero exponent corresponds to a step, or abrupt change in a function. In one dimension, a Lipschitz exponent of -1 is associated with a Dirac delta point source. In two dimensions, a Lipschitz exponent of -1 is indicative of a line source or linearly distributed concentration of some sort. An exponent of -2 corresponds to a point-like concentration or source in two dimensions. Thus, the Lipschitz exponent is similar in spirit to the 'structural index' encountered in the Euler methods. We will clarify the relationship between the two much later in our conclusions.

Let us briefly review the relationship between Lipschitz regularity and scale dependence of the wavelet transform as it applies to 1-D signals. The subsequent analysis is based on the following theorem, whose demonstration can be found in Holschneider & Tchamitchian (1989).

Theorem 1. Let $0 < \alpha \leq 1$. A function $f(x)$ is Lipschitz α over the interval (a, b) if and only if there exists a constant K such that, for all $x \in (a, b)$, the wavelet transform satisfies

$$|W[f](s, x)| \leq K s^{\alpha}\,. \tag{2.6b}$$

In Appendix A, we demonstrate the forward implication (2.6a implies 2.6b) which is the main basis for the following work. We also extend the forward implication to higher dimensions and more general singularities—in particular, logarithmic singularities. An indication of the treatment of negative exponents, and an outline of the machinery that applies to the extension to irregular sets is also deferred to Appendix A. The reverse implication (2.6b implies 2.6a) has been proved in the stated reference for the case of compactly supported wavelets. We have yet to apply the inverse transform in Appendix B to extend this proof to the wavelets that we will be using. (The inverse transform may not be entirely necessary to this endeavour.)

Locally continuous features are common in signals arising from physical experiments. They are characterized by positive Lipschitz exponents. Consequently, their wavelet coefficients decrease as the scale decreases. In the signal analysis literature, these continuous features are usually *modelled* as the convolution of singularities with Gaussian kernels. The Gaussian is, in turn, interpreted as the Green's function of the heat equation, so that the variance σ^2 becomes a measure of the time required to obtain a smoothed feature from some hypothetical causative singularity. In this way, the signal may be *modelled* as a linear superposition of various singularities, each convolved with its own (unique) Gaussian; that is, by a heat diffusion *approximation.* An estimation of the σ^2 of each (hypothetical) Gaussian and the Lipschitz exponent of the underlying singularity can then be calculated by studying the multiscale edges derived from Gaussian scaling functions. This idea is the basis of much of the work in signal and image processing in this area.

The foregoing analysis suggests that a useful signal model could also be obtained by replacing the Green's function of the heat diffusion problem with the Green's function of some potential field theory. In this case, features in the signal would be related to density/susceptibility sources at different distances (depths) below the measurement level. This approach is particularly suited to our problem since potential field maps *are* the result of such physical processes. Consequently, such a signal model would respect the true physics. The obvious conclusion is that we should also use the potential field's Green's function to construct the analysing wavelets, the hope being that the results relating to the Gaussian wavelets generalize.

2.4 Extension to potential field analysis

2.4.1 The basic equations

Consider the gravitational field. (The extension to magnetic fields is discussed later.) The gravitational potential at a point

(x, y, z_0) due to a source distribution $\rho(x, y, z)$ can be written as

$$V(x, y, z_0)$$
$$= -G \int_{\mathscr{R}^2} dx'\,dy' \int_{-\infty}^{0} \frac{\rho(x', y', x')\,dz'}{[(x-x')^2 + (y-y')^2 + (x_0 - z')^2]^{1/2}}, \tag{2.7}$$

where G is the gravitational constant and z_0 is the level at which the measurements have been taken. We also suppose that $\rho(x, y, z) = 0$ for $z > 0$. (That is, z increases upwards.) It follows that the magnitude of vertical acceleration is

$$f_{z_0}(x, y) = -g_z = V_{,3} = G \int_{\mathscr{R}^2} dx'\,dy'$$
$$\times \int_{-\infty}^{0} \frac{\rho(x', y', z')(z_0 - z')\,dz'}{[(x-x')^2 + (y-y')^2 + (z_0 - z')^2]^{3/2}}. \tag{2.8}$$

The expression

$$K(x, y, z) = \frac{z}{(x^2 + y^2 + z^2)^{3/2}} \tag{2.9}$$

is the Green's function for the vertical acceleration, and direct calculation shows that

$$\int_{\mathscr{R}^2} K(x, y, z)\,dx\,dy = 2\pi \tag{2.10}$$

is independent of z. Hence the integral of the function

$$\gamma_z(x, y) = \frac{1}{2\pi} K(x, y, z) \tag{2.11}$$

is unity for all $z > 0$. Continuing this line of reasoning, we find that

$$\begin{aligned} \gamma_{sz}(x, y) &= \frac{1}{2\pi} \frac{sz}{[x^2 + y^2 + (sz)^2]^{3/2}} \\ &= s^{-2} \gamma_z(x/s, y/s), \end{aligned} \tag{2.12}$$

which is the 2-D analogue of eq. (2.1). Consequently, (2.11) can be used to define a smoothing function, and hence a wavelet analysis of 2-D images.

From eqs (2.8,9) and (2.11), we see that the vertical acceleration measured at a level z_0 is given by

$$f_{z_0}(x, y) = 2\pi G \int_{-\infty}^{0} \rho(x, y, z')^* \gamma_{z_0 - z'}(x, y)\,dz', \tag{2.13}$$

where the symbol * now represents convolution on both x and y. In view of eqs (2.1)–(2.5), eqs (2.12) and (2.13) are more than a little suggestive of a relationship to a wavelet transform.

2.4.2 Upward continuation

Define the 2-D Fourier transform by

$$\hat{f}(\mathbf{k}) = \mathscr{I}[f](\mathbf{k}) = \int_{\mathscr{R}^2} f(\mathbf{x}) \exp(-2\pi i \mathbf{k} \cdot \mathbf{x})\,d\mathbf{x}, \tag{2.14}$$

where $\mathbf{x}$ is a 2-D vector and $\mathbf{k}$ is the 2-D phase vector. (Note that $\mathbf{k}$ is in cycles per unit length.) Then the Fourier transform of the Green's function for the vertical acceleration (for example Blakely 1995) yields

$$\hat{\gamma}_z(\mathbf{k}) = \exp(-2\pi \|\mathbf{k}\| z), \quad z > 0. \tag{2.15}$$

By the use of the convolution theorem, the Fourier transform of the gravity measurements can then be written as

$$\hat{f}_{z_0}(\mathbf{k}) = 2\pi G \int_{-\infty}^{0} \hat{\rho}(\mathbf{k}, z) \hat{\gamma}_{z_0 - z'}(\mathbf{k})\,dz'. \tag{2.16}$$

For a certain height $z > z_0$ we can express the vertical acceleration via the upward continuation relation:

$$\begin{aligned} \hat{f}_z(\mathbf{k}) &= 2\pi G \int_{-\infty}^{0} \hat{\rho}(\mathbf{k}, z') \exp[-2\pi \|\mathbf{k}\| (z - z')]\,dz' \\ &= 2\pi G \int_{-\infty}^{0} \hat{\rho}(\mathbf{k}, z') \exp\{-2\pi \|\mathbf{k}\| [(z_0 - z') + (z - z_0)]\}\,dz' \\ &= \hat{f}_{z_0}(\mathbf{k}) \hat{\gamma}_{z - z_0}(\mathbf{k}). \end{aligned} \tag{2.17a}$$

Thus

$$f_z(\mathbf{x}) = [f_{z_0} {}^* \gamma_{z - z_0}](\mathbf{x}), \quad z > z_0, \tag{2.17b}$$

which shows that the field continued to the level z is equal to the field at the measurement level z_0 convolved with the smoothing function for the proper height difference. The right-hand side of eq. (2.17b) is directly comparable to the term $[f^*\theta_s](x)$ in (2.5).

2.4.3 Wavelet transform

Our final task is to find a natural scale parameter s which leads to a convenient choice of the z values corresponding to $s = 0$ and $s = 1$. To this end, let $f_0(x, y)$ be the vertical acceleration at $z = 0$. Also, let the measurement height be z_0, and define the smoothing function at scale $s = 1$ by $\theta(x, y) = \gamma_{z_0}(x, y)$. From this, we construct the smoothing function for scale s from eq. (2.12); that is,

$$\theta_s(x, y) = \gamma_{sz_0}(x, y) = s^{-2} \theta(x/s, y/s). \tag{2.18}$$

The derivatives of the smoothing function with respect to x and y define the first-order wavelets, namely

$$\psi^1 = \theta_{,1} = D_x \theta \quad \text{and} \quad \psi^2 = \theta_{,2} = D_y \theta. \tag{2.19}$$

It is then easily verified that, for $s > 0$,

$$\psi_s^i(x, y) = s^{-2} \psi^i(x/s, y/s) \tag{2.20}$$

form a set of self-consistent dilation equations for these first-order wavelets.

It follows that the x-component of the 2-D wavelet transform of $f_0(x, y)$ is given by

$$\begin{aligned} W^1[f_0](s, x, y) &= [f_0 {}^* \psi_s^1](x, y) \\ &= f_0 {}^* s D_x \theta_s \\ &= s D_x [f_0 {}^* \theta_s] \\ &= s D_x [f_0 {}^* \gamma_{sz_0}] \\ &= (z/z_0) D_x f_z(x, y), \end{aligned} \tag{2.21a}$$

where $s = z/z_0$. Similarly, the y-component of the 2-D transform is

$$W^2[f_0](s, x, y) = (z/z_0) D_y f_z(x, y), \tag{2.21b}$$

and hence the vector 2-D wavelet transform is

$$\mathbf{W}[f_0](s, \mathbf{x}) = (z/z_0) \nabla f_z(\mathbf{x}), \tag{2.22a}$$

where ∇ denotes the 2-D gradient. This equation shows that the chosen wavelet transform of $f_0(x, y)$ at a certain scale $s = z/z_0$ can be obtained from measurements at a level z_0 by:

(1) upward continuing to the level $z = sz_0$;
(2) taking the 2-D horizontal gradient;
(3) multiplying by the factor s.

Eq. (2.22a) is the crux of the connection between wavelet transform theory and potential field theory, especially as it relates to gravity. It is also of more than passing interest that this forges a correspondence between the horizontal gravity gradients, and a wavelet transform of an 'effective planar source'.

Continuing on, we define the modulus of the transform as the scalar value

$$M[f_0](s, \mathbf{x}) = \|\mathbf{W}[f_0](s, \mathbf{x})\|\,, \tag{2.22b}$$

where the norm is the 2-D Euclidean distance. Then theorem 1 can be formulated as follows for the 2-D case (see Appendix A).

If $\rho(x, y, z')$ [as a function of (x, y)] is $\mathrm{LUR}(\rho) = \alpha$ at $z' < 0$, and $f_0^\delta(x, y)$ is the vertical acceleration generated by $\rho(x, y, z)\delta(z - z')$, then

$$M[f_0^\delta](s, x, y) \le K(s + s')^{\alpha - 1} s\,, \tag{2.23}$$

where $s = z/z_0$ and $s' = -z'/z_0$.

Note that $f_0^\delta(x, y)$ is the field at $z = 0$ generated by the concentrated planar source $\rho(x, y, z)\delta(z - z')$ and *not* the field generated by the whole of $\rho(x, y, z)$. The full field $f_0(x, y)$ is a superposition of such elementary contributions from all the planes $z' \le 0$, and we shall investigate some of the effects of this in due course. A general study of the relationship between the 3-D Lipschitz regularity of the source and the scaling of this 2-D wavelet transform is a matter of some delicacy. We have chosen instead to conclude as much useful information as we can by considering the singular behaviour of the *field* at various depths, rather than tackling the source behaviour itself. Since there is often a close correspondence between the singularity in the field on a plane passing through a source singularity and the source singularity itself, this is often a successful strategy. Another view is that we seek to model the source as a set of effective planar sources, and attempt to locate the singularities in these effective sources.

The right-hand side of (2.23) is a function of s, s' and α. By analysing the variation of the wavelet transform as a function of s, we can obtain estimates of s' and α. These estimates lead to the type of feature in the source (point source, line source, discontinuous jump or smooth variation), encoded in α, and the depth z' of the feature, encoded in s'.

In particular, it is the edges of (2.22b) (maxima in x and y along gradient streamlines) that are chosen to be the points at which the wavelet transform is studied. The s dependence of eq. (2.22b) at the edges amounts to the scale dependence of the 'sharpness' of the multiscale edge. We strongly recommend consultation of Mallat & Zhong (1992) for a more complete exposition of these matters.

Of course, the simple case of a planar source leading to (2.23) can act only as a guide to the analysis of the actual situation described by a continuous superposition of planar sources [as in eqs (2.13) and (2.16) for $z_0 = 0$]. Much of the following work is devoted to understanding and dealing with this difficulty by way of explicit examples of great relevance to the interpretation of geophysical data. We deliberately avoid a general mathematical discussion, as it would lead us quickly far from the matters immediately to hand. However, in the interest of self-containment, we do present some introductory material in Appendix A.

3 APPLICATIONS

Edge detection, via multiscale wavelet decomposition, and generalizations of eq. (2.23) offer a very powerful tool for at least two major analyses of potential field data: first, estimation of the depths and types of singularities in the source distribution; second, automatic generation of 'worm diagrams' or skeletonizations. These applications are described in the following sections.

3.1 2-D examples

Eq. (2.8) gives the general free-air formula for the vertical gravitational attraction in three dimensions. The 2-D case is commonly obtained by assuming that the distribution of density is independent of y; that is, the density distribution extends indefinitely perpendicular to the 2-D vertical section under analysis. There is a minor difficulty with the divergence of the energy (2.7) in some circumstances, but this is a harmless divergence typical of an extensive quantity of an infinite system. Energy differences (intrinsic quantities) generally remain finite and well defined.

Integration over y' in eq. (2.8) yields

$$f_{z_0} = -2G \int_{-\infty}^{\infty} dx' \int_{-\infty}^{0} \frac{\rho(x', z')(z_0 - z')\,dz'}{(x - x')^2 + (z_0 - z')^2}\,. \tag{3.1}$$

According to eq. (2.21b), the wavelet transform $\mathbf{W}[f_0](s, x, y)$ has a zero y-component and is independent of y. In this case, f_0 effectively becomes a 1-D signal, and we may take the wavelet transform as

$$W[f_0](s, x) = -2Gs \int_{-\infty}^{\infty} dx' \int_{-\infty}^{0} \rho(x', z') D_x\left[\frac{(s - z')\,dz'}{(x - x')^2 + (s - z')^2}\right]. \tag{3.2}$$

In (3.2) we have chosen our units so that $s = z$, and will continue to do so. (That is, we choose units of length so that $z_0 = 1$.) Eq. (3.2) can also be written either as

$$W[f_0](s, x) = -2Gs \int_{-\infty}^{\infty} dx' \int_{-\infty}^{0} \rho_{,1}(x', z') \frac{(s - z')\,dz'}{(x - x')^2 + (s - z')^2} \tag{3.3}$$

or as

$$W[f_0](s, x) = 4Gs \int_{-\infty}^{\infty} dx' \int_{-\infty}^{0} \rho(x', z') \frac{(s - z')(x - x')\,dz'}{[(x - x')^2 + (s - z')^2]^2}\,. \tag{3.4}$$

3.1.1 2-D block

Consider the case of a 2-D rectangular block, of constant density, whose *density contrast* with the surrounding material is unity. The configuration is sketched in Fig. 1.

The 2-D section of the block in the x–z plane is between $x = 0$ and $x = x_{\max}$ and between $z = z_l$ and $z = 0$. Employing eq. (3.3) for the wavelet transform of the gravity anomaly we

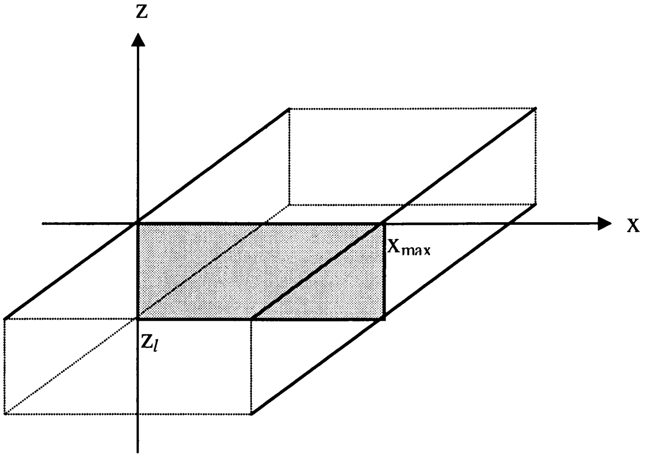

Figure 1. 2-D rectangular prism extending indefinitely in the *y*-direction.

find

$$
\begin{aligned}
W[f_0](s,x) \\
&= 2Gs \int_{-\infty}^{\infty} dx' \int_{z_l}^{0} [\delta(x') - \delta(x' - x_{\max})] \frac{(s - z')\,dz'}{(x - x')^2 + (s - z')^2} \\
&= 2Gs \int_{z_l}^{0} (s - z') \left(\frac{1}{x^2 + (s - z')^2} - \frac{1}{(x - x_{\max})^2 + (s - z')^2} \right) dz', \qquad (3.5)
\end{aligned}
$$

where $\delta(x)$ is the Dirac delta function. Notice that we are interested in points of sharp variation, which correspond to local extremes of the wavelet transform, independent of the sign. Accordingly, we shall often ignore the sign of the amplitude of the maxima.

Performing the integration (3.5) yields

$$
W[f_0](s,x) = Gs \log \left\{ \frac{(x^2 + (s - z_l)^2)[(x - x_{\max})^2 + s^2]}{(x^2 + s^2)[(x - x_{\max})^2 + (s - z_l)^2]} \right\}. \qquad (3.6a)
$$

If the block extends from z_l to z_d (the top of the block is buried at depth z_d) then (3.6a) is modified to

$$
\begin{aligned}
W[f_0](s,x) &= Gs \log \left\{ \frac{(x^2 + (s - z_l)^2)[(x - x_{\max})^2 + (s - z_d)^2]}{(x^2 + (s - z_d)^2)[(x - x_{\max})^2 + (s - z_l)^2]} \right\} \\
&= 2Gs[\log(d_{0l} d_{md}) - \log(d_{0d} d_{ml})], \qquad (3.6b)
\end{aligned}
$$

where the $d_{\xi\zeta}$ are the distances from the point x to the corners of the rectangular cross-section. [For example, compare with the results for (infinite) parallelepipeds in Telford *et al.* (1976), pp. 72–74.]

The location and behaviour of the wavelet maxima define the multiscale edges. Their location can be obtained by calculating the zeros of the first derivative of (3.6a) with respect to x. The solutions have the form

$$
x = \{6x_{\max} \pm \sqrt{12x_{\max}^2 - 24[s^2 + (s - z_l)^2 - A]}\}/12, \qquad (3.7)
$$

where

$$
A = \sqrt{12[s(s - z_l)]^2 + [2s(s - z_l) + x_{\max}^2 + z_l^2]^2}. \qquad (3.8)
$$

This equation tells us that the wavelet maxima are displaced with respect to the border of the block, the displacement depending on the values of z_l, $x_{\max}$ and s.

Fig. 2(a) shows that the displacement of the wavelet maxima relative to the block border increases almost linearly with s for narrow blocks, whereas this departure gets smaller as the width of the block increases. Similarly, Fig. 2(b) shows that the displacement is larger for thick blocks and increases with increasing thickness. This confirms previous results presented by Grauch & Cordell (1987). By substituting the x-value (3.7) in (3.6a), the variation of the amplitude of the wavelet maxima ($W_{\max}$) at different scales can be obtained.

Plots of the evolution of the wavelet maxima at different scales for different values of z_l are presented in Fig. 3. The plots show a fast increase in $W_{\max}$ for small s up to a maximum,

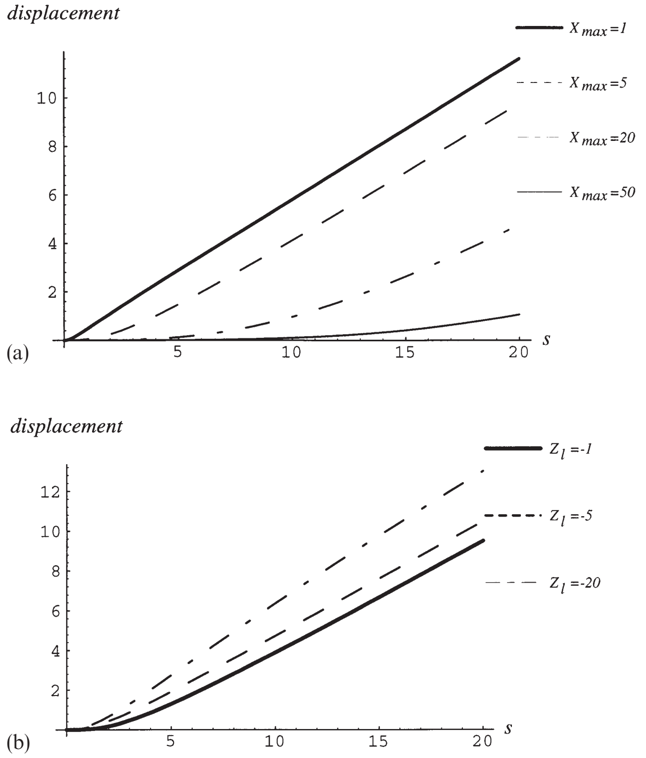

Figure 2. Displacement of the wavelet maxima (edges) relative to the border of the causative block: (a) for various values of $x_{\max}$; (b) for various z_l.

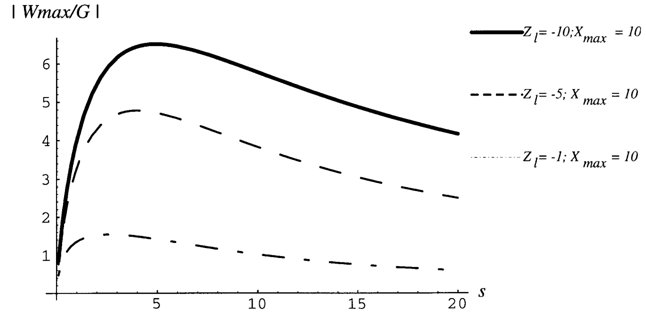

Figure 3. Evolution of the amplitude of the wavelet maxima ($W_{\max}$) at different levels above the top of the block for various values of z_l. The scale of the maximum in the curve increases for increasingly thick blocks.

after which $W_{\max}$ decreases slowly. The information contained in the location and amplitude of the wavelet maxima can be used in an inversion procedure to determine the characteristics and location of the causative edge or contrast in $\rho(x, y, z)$. An example of this approach is presented in Section 4.

3.1.2 *Tilted dyke*

Another common structure used in 2-D geological modelling is the infinitely thin dyke, as in Fig 4. The equation for the wavelet transform can be derived from (3.4) as

$$
\begin{aligned}
W[f_0](s, x) &\\
&= 4Gs \int_0^{\infty} dx' \int_{-\infty}^{0} \delta(-hx' - z') \frac{(s - z')(x - x')\, dz'}{[(x - x')^2 + (s - z')^2]^2} \\
&= \frac{2Gs(hx - s)}{(1 + h^2)(s^2 + x^2)} . \qquad (3.9)
\end{aligned}
$$

Following the same steps as in the previous section, we look for the locations at which the wavelet maxima occur. The zeros of the derivative of (3.9) with respect to x are

$$x = \frac{s(1 \pm \sqrt{1 + h^2})}{h} , \qquad (3.10)$$

which shows that the locations of the wavelet maxima vary linearly with s. By substituting (3.10) in (3.9) we obtain the amplitudes of the wavelet maxima:

$$W_{\max} = \frac{Gh^2}{\sqrt{1 + h^2}[\sqrt{1 + h^2} \pm (1 + h^2)]} . \qquad (3.11)$$

This value depends only on the inclination of the dyke and is constant at each level above the surface (that is, constant in s). This behaviour makes the anomaly from a tilted dyke easily distinguishable from that of the 2-D block. Indeed, eqs (2.6) and (2.23) indicate that the variation of $W_{\max}$ with s is the primary distinguishing feature determining the nature of the singularity in $\rho(x, y, z)$. Before discussing this matter in greater detail, let us consider one more 2-D case.

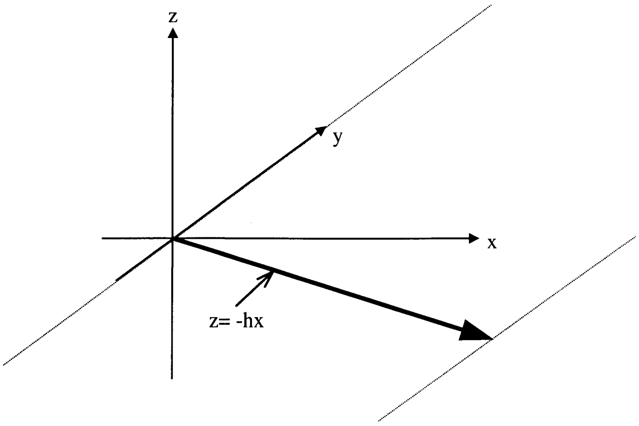

Figure 4. Tilted dyke extending indefinitely in the positive-x and the y-directions. The inclination of the dyke is given by the equation $z = -hx$.

3.1.3 *Line source*

Long, thin bodies can sometimes be modelled as line sources. Here we study the wavelet transform of a line source buried at depth z_d, as sketched in Fig. 5. In this case the result can easily be read from eq. (3.4):

$$W[f_0](s, x) = -4Gs \frac{x(s - z_d)}{[x^2 + (s - z_d)^2]^2} . \qquad (3.12)$$

Again, the locations of the wavelet maxima are found by differentiating (3.13) with respect to x and equating to zero, so that

$$x = \pm(s - z_d)/\sqrt{3} , \qquad (3.13)$$

and the amplitude of the wavelet maxima is given by

$$W_{\max} = \pm \frac{3\sqrt{3}}{4} \frac{Gs}{(s - z_d)^2} . \qquad (3.14)$$

3.1.4 *Summary of 2-D results*

At this point it is worth reviewing the results of these three examples in the light of Theorem 1 and the related equations (2.6a,b) and (2.23). First, we note that the two multiscale edges described by eqs (3.13) and (3.14) (from the line source) fit into the category of $\alpha = -1$ in eq. (2.23). This is precisely the value expected from the nature of the singularity in the density field, being of Lipschitz uniform regularity -1. Moreover, any such concentrated source can be expected to yield two multiscale edges, corresponding to the left and right 'edges' of the singularity. We have termed such pairing of multiscale edges 'dipole edges'. Such edge pairings are often associated with negative Lipschitz exponents when the first-order wavelets ψ_s of the form (2.2) and (2.19) are used.

Eqs (3.10) and (3.11) (sheet singularity) are typical of $\alpha = 0$ in eq. (2.23). Indeed, if one takes $h \to 0$, one finds that the left multiscale edge tends to $x = 0$, while its amplitude remains non-zero and independent of s. The weaker multiscale edge to the right of $x = 0$ becomes weaker and moves off to infinity as

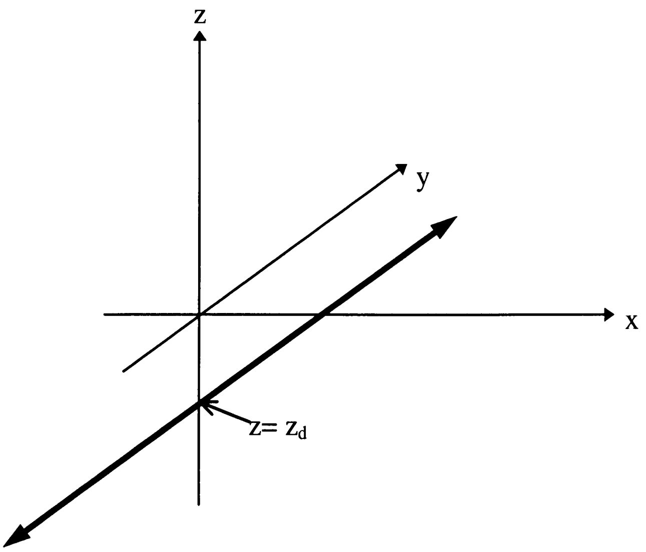

Figure 5. Line source at depth z_d, perpendicular to the x–z plane, extending indefinitely in the y-direction.

$h \to 0$. The $h = 0$ case is precisely a step function of Lipschitz uniform regularity 0 on the plane $z = 0$. Hence the behaviour as $h \to 0$ is exactly as predicted by (2.23). What our dyke example has achieved is to show how the behaviour (2.23) is modified for a tilted planar source. We see the appearance of an asymmetric dipole edge response, which is indicative of the dip of the planar source. Nevertheless, the stronger multiscale edge retains its $\alpha = 0$ character.

Finally, we come to the most complex case, eqs (3.6)–(3.8) (density jump singularity). This case does not fit into the Lipschitz classification scheme, and so (in Appendix A), we make a modest extension to the notion of Lipschitz exponent. Consider the case of eq. (3.6a) (that is, $z_d = 0$). In this case, the actual vertical acceleration at $z = 0$ (that is, $f_0(x)$) has infinite slope at $x = 0$. This divergence in the slope is logarithmic. Consequently, $f_0(x)$ is Lipschitz $(1 - \varepsilon)$ for all small $\varepsilon > 0$. However, $f_0(x)$ is *not* Lipschitz 1. The trick for dealing with this situation is to modify (2.6a) by setting $\alpha = 1$ and introducing a factor $|\log|x_0 - x_1||$ on the right-hand side, which we do in Appendix A. Setting $x = 0$ in (3.6a) then gives the satisfying result

$$W[f_0](s, 0) \sim K|s \log s| \quad \text{as } s \to 0 ,$$

as predicted in Appendix A. The behaviour as $s \to \infty$ is a matter of some delicacy. As a result of some remarkable cancellations, (3.6a) tends to zero (at $x = 0$) as $s \to \infty$, whenever $x_{\max}$ is finite. However, the $s \to \infty$ limit is non-zero if $x_{\max}$ is infinite, that limit being determined by z_l.

3.2 3-D examples

We can now extend the previous results to 3-D configurations. From eq. (2.8), the vertical gravitational acceleration can be written as

$$f_z(x, y) = -G \frac{\partial}{\partial z} \int_{\mathscr{R}^3} \frac{\rho(\boldsymbol{x}')}{\|\boldsymbol{x} - \boldsymbol{x}'\|} v^{(3)}(d\boldsymbol{x}') \tag{3.15}$$

where $v^{(3)}$ is a 3-D volume measure and bold italic symbols represent 3-D vectors. We know that the derivatives with respect to x and y give the wavelet transform, so let us consider the 3-D gradient operator $\nabla_{\boldsymbol{x}}$ and select the desired derivatives by projection. Now

$$\begin{aligned} \nabla_{\boldsymbol{x}} f_z(x, y) &= -G \frac{\partial}{\partial z} \int_{\mathscr{R}^3} \rho(\boldsymbol{x}') \nabla_{\boldsymbol{x}} \frac{1}{\|\boldsymbol{x} - \boldsymbol{x}'\|} v^{(3)}(d\boldsymbol{x}') \\ &= -G \frac{\partial}{\partial z} \int_{\mathscr{R}^3} \nabla_{\boldsymbol{x}'} \rho(\boldsymbol{x}') \frac{1}{\|\boldsymbol{x} - \boldsymbol{x}'\|} v^{(3)}(d\boldsymbol{x}') . \end{aligned} \tag{3.16}$$

For convenience, suppose ρ can take only the value 0 or 1, and define $\Omega = \{\boldsymbol{x}: \rho(\boldsymbol{x}) = 1\}$. Furthermore, let $\partial\Omega$ denote the boundary of Ω, and let it be a regular, rectifiable 2-D set.

Next we define a local parametrization $\partial\Omega = \{\boldsymbol{x}: \boldsymbol{x} = \boldsymbol{x}(\xi, \eta, \zeta_0)\}$, where ζ is the path length perpendicular to $\partial\Omega$, and $\hat{\boldsymbol{e}}_\xi, \hat{\boldsymbol{e}}_\eta, \hat{\boldsymbol{e}}_\zeta$ the corresponding unit normals in the ξ, η, ζ coordinate system (see Fig. 6). We are free to choose the scale of ζ so that $\nabla_{\boldsymbol{x}} \rho = \delta(\zeta - \zeta_0) \hat{\boldsymbol{e}}_\zeta$. Since we have assumed unit density contrast, this identifies ζ as the path length normal to $\partial\Omega$, as stated earlier. Eq. (3.16) becomes

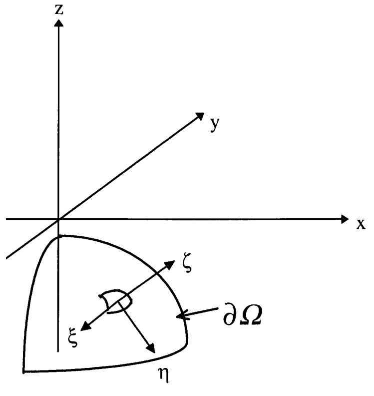

Figure 6. Singular surface representing the boundary between regions of different density.

$$\begin{aligned} \nabla_{\boldsymbol{x}} f_z(x, y) &= -G \frac{\partial}{\partial z} \int_{\mathscr{R}^3} \delta(\zeta - \zeta_0) \hat{\boldsymbol{e}}_\zeta \frac{v^{(3)}(d\boldsymbol{x}')}{\|\boldsymbol{x} - \boldsymbol{x}'\|} \\ &= -G \frac{\partial}{\partial z} \int_{\mathscr{R}^3} \frac{\delta(\zeta - \zeta_0) \hat{\boldsymbol{e}}_\zeta}{\|\boldsymbol{x} - \boldsymbol{x}'(\xi, \eta, \zeta)\|} J(\xi, \eta, \zeta)\, d\xi\, d\eta\, d\zeta \\ &= -G \frac{\partial}{\partial z} \int_{\delta\Omega} \frac{1}{\|\boldsymbol{x} - \boldsymbol{x}'(\xi, \eta, \zeta_0)\|} J(\xi, \eta, \zeta_0)\, d\xi\, d\eta\, \hat{\boldsymbol{e}}_{\zeta_0} . \end{aligned}$$

Now ζ is, locally at $\partial\Omega$, a path length, so that the Jacobian $J(\xi, \eta, \zeta_0)$ is the (area) ratio of $\|d\xi \times d\eta\|$ to $d\xi d\eta$ in 3-D space. Thus $J(\xi, \eta, \zeta_0)\, d\xi\, d\eta$ is the 3-D metric area of an infinitesimal element of $\partial\Omega$. Moreover, $\hat{\boldsymbol{e}}_{\zeta_0}$ is the unit normal to this area. So

$$\nabla_{\boldsymbol{x}} f_z(x, y) = -G \frac{\partial}{\partial z} \int_{\delta\Omega} \frac{dA'}{\|\boldsymbol{x} - \boldsymbol{x}'\|} = G \int_{\delta\Omega} \frac{(z - z')\, d\boldsymbol{A}'}{\|\boldsymbol{x} - \boldsymbol{x}'\|^3} .$$

Scaled by the density contrast, this equation gives the vector wavelet transform for piecewise-constant density fields in terms of the singular surfaces. This expression is geometrical (coordinate-free) involving normals and areas, independent of the parametrization.

By defining $\boldsymbol{O}_z \boldsymbol{v} = \boldsymbol{v} - (\boldsymbol{v} \cdot \hat{\boldsymbol{e}}_z) \hat{\boldsymbol{e}}_z$, the projection perpendicular to the z-axis, the wavelet transform in the 3-D gravitational case can be written in one of the forms

$$\mathbf{W}[f_0](s, x, y) = G s \boldsymbol{O}_z \int_{\delta\Omega} \frac{(s - z')\, d\boldsymbol{A}'}{[(x - x')^2 + (y - y')^2 + (s - z')^2]^{3/2}} , \tag{3.17a}$$

which is convenient for piecewise-constant densities, or

$$\mathbf{W}[f_0](s, x, y) = G s \int_{\mathscr{R}^3} \rho(\boldsymbol{x}') \nabla \left[\frac{(z - z')}{\|\boldsymbol{x} - \boldsymbol{x}'\|^3} \right]_{z=s} v^{(3)}(d\boldsymbol{x}') \tag{3.17b}$$

or

$$\mathbf{W}[f_0](s, x, y) = G s \left[\frac{\partial}{\partial z} \int_{\mathscr{R}^3} \rho(\boldsymbol{x}') \nabla \frac{1}{\|\boldsymbol{x} - \boldsymbol{x}'\|} v^{(3)}(d\boldsymbol{x}') \right]_{z=s} , \tag{3.17c}$$

which are appropriate for sheet, line and point sources.

3.2.1 3-D block of infinite strike

As an application of eq. (3.17a), we derive the vertical acceleration due to a prismatic body between $z=0$ and $z=z_l$, of infinite strike in the y-direction. The left border dips with slope $-h$, reaching the surface at $x=0$. The right border is at $x=\infty$ (see Fig. 7). Once again, we assume a unit density contrast.

Applying eq. (3.17a), we first note that the top and bottom surfaces do not contribute, since they are projected away by $\boldsymbol{O}_z$. This leaves the contribution from the dipping plane. We can parametrize this surface with y' and l (see Fig. 7). Given the dip of the plane $(-h)$, the element of surface can (dropping the z-component) be written as

$$dA' = \frac{dy'\,dl}{\sqrt{1+\eta^2}}\begin{bmatrix}1\\0\end{bmatrix}, \quad \text{with } \eta = h^{-1}. \tag{3.18}$$

Accordingly, (3.17a) becomes

$$\mathbf{W}[f_0](s,x,y)$$
$$= 2Gs\int_0^{l_{\max}} \frac{[s-z'(l)]}{[x-x'(l)]^2+[s-z'(l)]^2}\begin{bmatrix}1/\sqrt{1+\eta^2}\\0\end{bmatrix} dl$$

(since z' and x' are independent of y'). Performing the substitution $dl=\sqrt{1+\eta^2}\,dz'$ then yields

$$W^1[f_0](s,x,y) = 2Gs\int_{z_l}^{0}\frac{(s-z')}{(x+\eta z')^2+(s-z')^2}dz',$$
$$W^2[f_0](s,x,y) = 0. \tag{3.19}$$

It can be verified that the x-component in (3.19) for $h=\infty$ (that is, for vertically dipping contact) is equal to the $x_{\max}=\infty$ case of eq. (3.5), as it should.

3.2.2 Sphere

Consider a sphere of radius a, and centre at $z=z_{\rm d}$. The vertical acceleration at (x, y, z) is

$$f_z(x,y) = -\frac{4\pi G\rho a^3}{3}\frac{z-z_{\rm d}}{[x^2+y^2+(z-z_{\rm d})^2]^{3/2}}, \tag{3.20}$$

where ρ is the density of the sphere and G the gravitational constant. As usual, we will take the density as unity, and suppose $z_{\rm d} < -a$.

The wavelet transform in x and y can be written as the derivative of the field multiplied by the scale factor. Thus

$$W^1[f_0](s,x,y) = 4\pi G a^3\frac{xs(s-z_{\rm d})}{[x^2+y^2+(s-z_{\rm d})^2]^{5/2}}. \tag{3.21a}$$

From rotational symmetry and eq. (2.22a) it follows that

$$M[f_0](s,x,y) = 4\pi G a^3\frac{s(s-z_{\rm d})\sqrt{x^2+y^2}}{[x^2+y^2+(s-z_{\rm d})^2]^{5/2}}, \tag{3.21b}$$

which defines the modulus of the wavelet transform. We concentrate our analysis on the plane $y=0$. The edges can be found by equating the first derivative of (3.21a) to zero, yielding the intersections of the multiscale edges with the $y=0$ plane,

$$x = \pm(s-z_{\rm d})/2. \tag{3.22}$$

This describes the location of the wavelet maxima at various levels above the centre of the sphere (Fig. 8). The amplitudes of the wavelet maxima are found to be

$$W_{\max} = \frac{64G\pi a^3}{25\sqrt{5}}\frac{s}{(s-z_{\rm d})^3}. \tag{3.23}$$

These results are valid for any vertical plane passing through the centre of the sphere, and accordingly describe the evolution of the wavelet maxima in three dimensions. One need only replace x in eq. (3.22) with the cylindrical radial coordinate to track the general 3-D multiscale edge.

There is a clear interpretation as $\alpha=-2$ in eq. (2.23), which fits perfectly within the classification scheme established previously for line, sheet and block sources provided one recalls that the sphere is equivalent to a point source at its centre. A note in passing is that, while the multiscale edges contain enough information to establish the depth and mass of the spherical body, there is no way of determining its size without knowledge of the density. This is a consequence of f_z being a function of ρa^3 only, and is not in any way caused by the wavelet representation.

4 INVERSION OF POTENTIAL FIELD DATA IN THE WAVELET DOMAIN

In this section we illustrate one way that the wavelet transform can be used for inversion purposes. To this end, a number of

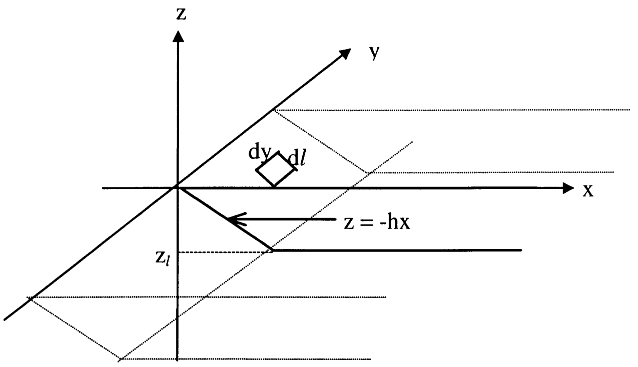

Figure 7. Prismatic body of thickness $z=z_l$, infinite strike in the y-direction and extending to infinity on the positive x-axis. The slope of the plane facing the negative x-direction is $-h$.

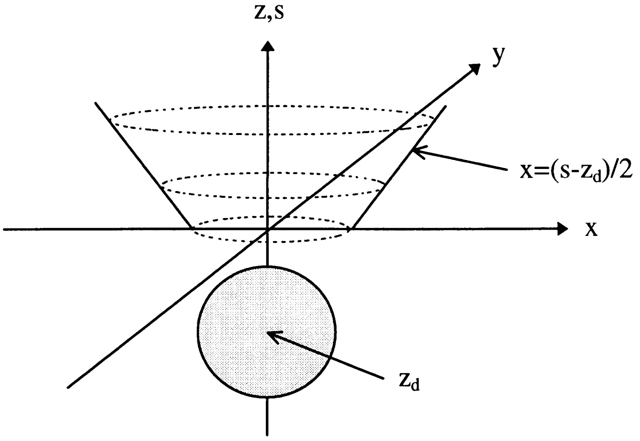

Figure 8. Location of the wavelet maxima at various levels above a sphere.

2-D synthetic examples of increasing complexity will be presented. Our objective is to use the information contained in the wavelet maxima to reconstruct the location and characteristics of the causative body.

Fig. 9 shows our first model, and corresponding data. It is a 2-D rectangular block, as in Section 3.1.1, buried to a depth of 10 m. In order to simulate a modestly realistic experiment, the horizontal location of the borders of the block do not lie on grid points at which the gravity profile is calculated, but are positioned midway between grid points. Also, the gravity profile was sampled at a spacing relatively coarse (1 m) compared with the width of the body. The relatively coarse spacing causes inaccuracy in the derivative calculation, while the non-coincidence with the grid results in errors in the location of the wavelet maxima. Fig. 10 shows the modulus of the resulting wavelet transform of the acceleration.

The amplitude of the wavelet maximum as a function of s is plotted in Fig. 11 (G is in units of metres, grams and seconds). The shape of the curve closely resembles the plot in Fig. 3. In fact, we can infer that the causative edge is a density contrast caused by the edge of a body of finite extent in the x-direction. The simplest assumption is a rectangular prism, and one might try this first.

Given the width, depth, thickness and density contrast of a prism, the amplitude of the wavelet maximum can be calculated as a function of s. Minimizing the squared error between this prediction and the data of Fig. 11 is a four-parameter minimization problem. Our first attempt to perform the inversion, using a local descent method, resulted in unacceptable sensitivity to the selection of the initial prism geometry, indicating that the minimization algorithm was being trapped by local minima of the error surface. Consequently, a global search method was employed, specifically, a genetic algorithm (GA). See, for example, Boschetti *et al.* (1996) for details of the particular implementation used.

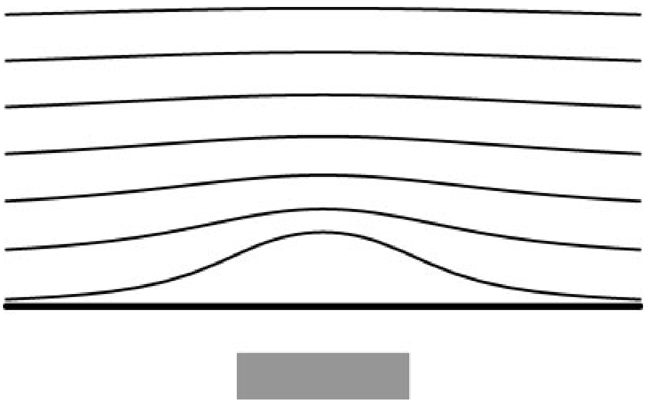

Figure 9. Data set employed in the first test. The block width is 30 m, and its thickness is 10 m. The depth to the top is 10 m, and the density contrast with the background is 0.3 g m^{-3}. Above the body we show the gravitational anomaly at various levels.

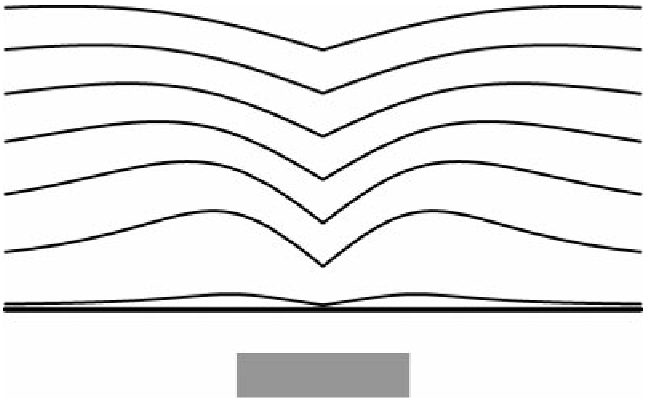

Figure 10. Modulus of the wavelet transform of the gravitational field caused by the body shown.

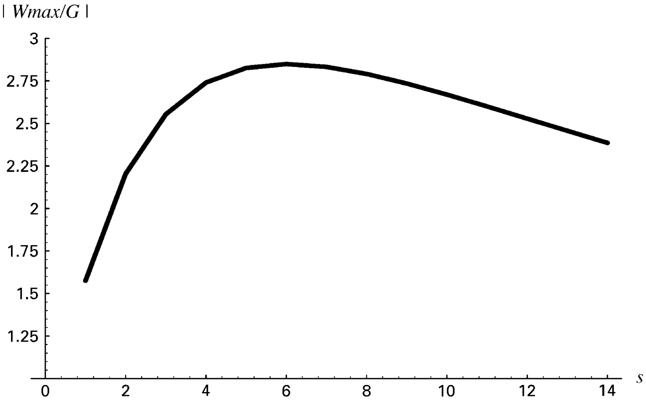

Figure 11. The amplitude of the multiscale edge as a function of scale s. The shape of the plot resembles that of Fig. 3, suggesting a 2-D rectangular prism as the source.

In order to assess the variability of the results obtained from stochastic global inversion techniques like our GA, the inversion was performed starting from 10 different, randomly chosen, initial prisms. Table 1 shows the results of the inversion. The best result together with the average and the standard deviation of the different runs are presented alongside the geometry of the prism that generated the data set.

The best solution (second column in Table 1) is a very good approximation of the true causative body, although there is no way we could know this in general. The average solution is the one of interest, as well as the variability of the solutions. We see that the results from the various GA runs differ quite substantially, especially in the estimation of the density contrast.

In our second test, the *locations* of the wavelet maxima, as well as their *amplitudes*, are used for the inversion, and the results presented in Table 2. It can be seen that the variability of the different GA runs is reduced and the estimation of the density contrast is more accurate. We also note that only one of the two multiscale edges was used for these inversions.

In the third test we increase the complexity of the experiment to simulate a slightly more realistic configuration. Here we model the presence of a second body, a tilted dyke, in the

Table 1. Results of the inversion of the synthetic data of Fig. 10 using *only* the wavelet maximum *amplitude* information. The standard deviation of the various GA runs is indicated in the right-most column.

Parameter	Synthetic	Best sol.	Av. sol.	Std dev.
Width (m)	30	29.5	29.1	2.0
Depth (m)	10	10.8	11.5	1.0
Thickness (m)	10	9.6	8.1	2.4
Density (g m^{-3})	0.3	0.32	0.40	0.10

Table 2. Results of the inversion for both the *amplitudes* and *locations* of the wavelet maxima, for the synthetic data set of Fig. 10.

Parameter	Synthetic	Best sol.	Av. sol.	Std dev.
Width (m)	30	29.5	29.7	0.3
Depth (m)	10	10.0	10.0	0.2
Thickness (m)	10	11.4	11.3	0.6
Density (g m^{-3})	0.3	0.28	0.27	0.01

vicinity of the block, as shown in Fig. 12. The presence of the dyke affects the wavelet transform at both edges of the block. The resulting inversion of the left block edge is shown in Table 3. We can see that the effect of the dyke is to give a moderately deeper and wider estimation of the causative block. However, the errors are still below 10 per cent, and accordingly the result of the inversion should be considered satisfactory.

5 WAVELET METHODS, INVERSION AND UNIQUENESS

The wavelet methods should not be viewed as some sort of magic way of extracting more information from a potential field map (image) than can be obtained from traditional methods. They are merely a different way of recovering the same information. The essence of what is happening is as follows.

The information about the type of causative edge in the source distribution is encoded in the shape of the graph of $f_{z_0}(x, y)$ regarded as a function of (x, y). Information about the shape of the graph of $f_{z_0}(x, y)$ is already commonly used to infer the characteristics of the causative body (for example Telford *et al.* 1976). However, the description of the shape of an anomaly in the graph of $f_{z_0}(x, y)$ first requires that the anomaly be isolated from adjacent ones, and then the parameters of its shape estimated. In complex data sets, this can be a difficult task to achieve; more so to automate.

By introducing the wavelet transform we do two things to change this situation:

(1) we introduce the vertical variation of the field into the picture; and

(2) we study the vertical variation of the horizontal derivatives.

Moreover, we specifically focus attention on the sharpness of the multiscale edges as a function of height.

The multiscale edges form a 2-D manifold that can be naturally embedded in the 3-D space above the image of the measured data. If each point of the multiscale edges is coloured according to the magnitude of the wavelet transform at that point, and transparency judiciously assigned, the appearance is not unlike an aurora. (We leave it to the reader to ponder this similarity in appearance.) It turns out that this information amounts to a different representation of the very same information as is contained in the shape of the graph of $f_{z_0}(x, y)$.

Points in favour of the multiscale edge representation of the information are these. A maximum, when perturbed by something slowly varying, tends to remain a maximum, and is therefore not unduly affected by other features, provided they are not too close. Moreover, the maximum is a property of a point (x, y). In contrast, the shape of an anomaly in the graph of $f_{z_0}(x, y)$ is a property of an extended region. Consequently, the determination of the location and amplitude of the wavelet maxima is an easy task to automate in comparison to the estimation of the shape of the graph of $f_{z_0}(x, y)$.

So far, our synthetic inversions seem acceptable when information on both the amplitude and location of the wavelet maxima is inverted. The quality of the results is only marginally affected by typical discretization inaccuracies, and there also seems to be some hope of disentangling multiple causative bodies, as discussed in the third experiment above. We also note that the estimation of the location and characteristics of the prismatic body were obtained by inverting the information from only one of two edges of the anomaly. Thus, in some (admittedly rare) circumstances, there is redundant information in the multiscale edges that could be exploited if one edge is disrupted.

The issue of the uniqueness of the inversion deserves particular attention and, in principle, we have the rigorous results of wavelet transform theory available to tell us what to expect. A Fourier analysis of the ambiguity domain in potential field inverse problems can be found in Mareschal (1985) (see also Boschetti *et al.* 1998), and the wavelet approach cannot be expected to ameliorate the fundamental problem demonstrated there. However, it may illuminate questions regarding the minimal constraints that render the inverse problem unique, especially constraints that are awkward to represent in the Fourier domain. Nevertheless, a more complete analysis will have to appear elsewhere. Here we will restrict attention to specific points raised in this paper.

The tests presented in the previous section show that the results from random initialization of the GA converge towards very similar solutions. This suggests that the problem is sometimes unique for piecewise-constant density fields. However, our example with the sphere demonstrates that this cannot be the case in general, even for simple parametrized shapes of constant density. This is not a problem with the multiscale edge approach, but rather a manifestation of the fact that any number of different spheres will yield the same potential field external to the sphere.

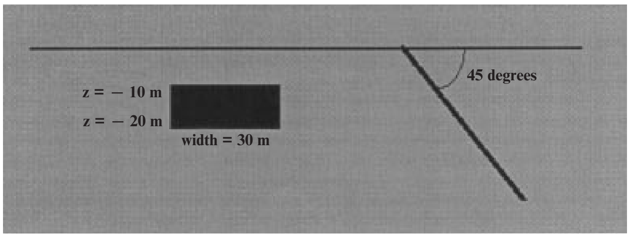

Figure 12. Model used in the third synthetic test. The experiment has been designed to test the effect of the presence of another causative body (in this case a dyke), in the vicinity of the block.

Table 3. Results of the inversion of both *amplitudes* and *locations* of wavelet maxima for the synthetic profile generated from Fig. 12.

Parameter	Synthetic	Best sol.	Av. sol.	Std dev.
Width (m)	30	32.7	32.7	0.2
Depth (m)	10	11.0	11.0	0.4
Thickness (m)	10	10.1	10.2	1.5
Density (g m^{-3})	0.3	0.30	0.31	0.05

Nevertheless, for Lipschitz exponents other than -2, we found unique inversions when the domain was constrained to have piecewise-constant density and a simple shape. Such assumptions are common to many inversion schemes (see Telford *et al.* 1976; Blakely 1995), and can be seen as *a priori* constraints on the causative body.

While the methodology of assuming some simply parametrized shape, and inverting for that shape, is an important approach, overemphasizing it misses what we feel is the main thrust of the method. The strength of the analysis is its ability to characterize and locate *singularities* in the density field, be they compact singular bodies of negative Lipschitz exponent, sheet-like sources of zero exponent, or geological contacts of 'one minus log' Lipschitz type.

By way of analogy, consider the Werner deconvolution approach (Werner 1953). Part of the analysis seeks some singular body (a sheet source) that explains a 'patch' of the potential field measurements. Applications of the method then collect point estimates of the sheet top (obtained from many patches) to piece together a view of the body causing the anomaly. In a sense, the wavelet method extends the Werner deconvolution method to a wider class of singularities, along the lines of a Euler deconvolution. Just as the singularities of these traditional methods can be used to infer the possible positions of causative bodies, so also can the point, line, sheet and contact singularities located by the wavelet analysis, and perhaps to better effect.

6 AUTOMATIC SKELETONIZATION OF REAL DATA

It has already been noted that most current analysis of potential field data is carried out visually by experienced geoscientists, and is directed towards the detection of geological contacts, faults and other features that are well characterized by singularities in the source distribution. This process is equivalent to edge detection, and such algorithms have been applied to potential field data for quite some time (see, for example, Blakely & Simpson 1986). Here we use the wavelet transform primarily as an edge detection method. The difference is that we construct multiscale edges that have the advantage of forming the basis for subsequent analysis of the depth and type of singularity.

For geophysical exploration purposes, magnetic maps are used more commonly than gravity. The greater relative variation of susceptibility and/or magnetization leads to greater contrast, and some economically interesting minerals have a strong magnetic response. When edge detection algorithms are applied to magnetic maps it is common practice to convert them first to pseudogravity (Blakely & Simpson 1986). There are two reasons for this. First, pseudogravity transformation involves reduction to the pole, which reduces directional anisotropy of the magnetic response to the source distribution. Second, transformation to pseudogravity leads to a single edge response for each source border, instead of the two edges present in the raw magnetic data. This difference is due to an extra differentiation that occurs in the Green's function for the total magnetic anomaly measurement (Telford *et al.* 1976), which is in turn a consequence of the dipolar nature of magnetic sources.

While the pseudogravity transformation includes a number of assumptions that are not generally valid, it allows a simple wavelet analysis on this type of data. A combination of upward continuation and the analytic signal represents an alternative approach, while the magnetostatic scalar potential and its relation to the 'monopole density' (divergence of the magnetization) is yet another possibility.

The automatic edge detection algorithm has been applied to the aeromagnetic data shown in Fig. 13. We see some strong magnetic anomalies emerging from a mostly flat background in which structure cannot be discriminated easily. Fig. 14 shows the edges detected by the algorithm at the finest scale. Not only are the major features in the data well defined by the edges, but the algorithm is also able to detect features in

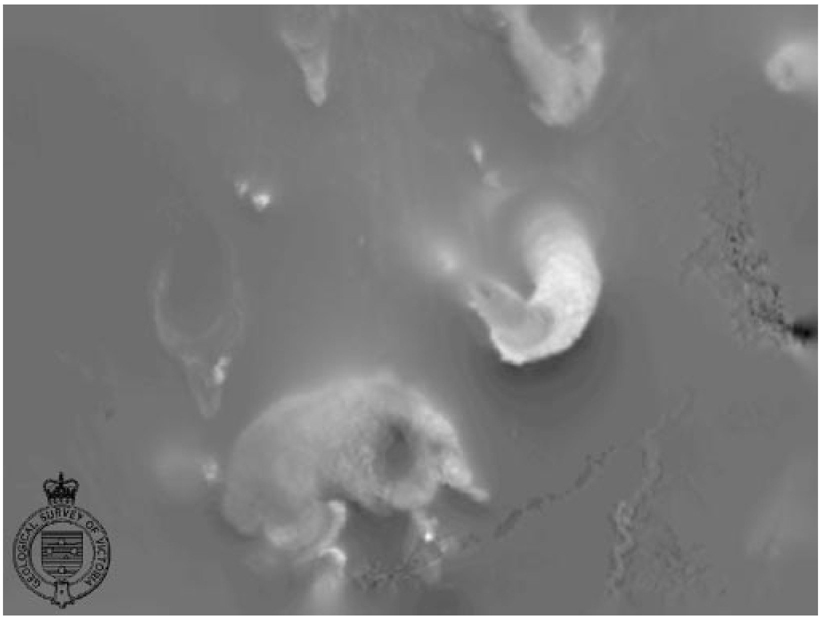

Figure 13. Greyscale interpolated aeromagnetic data for a region near St Arnaud, Victoria, Australia.

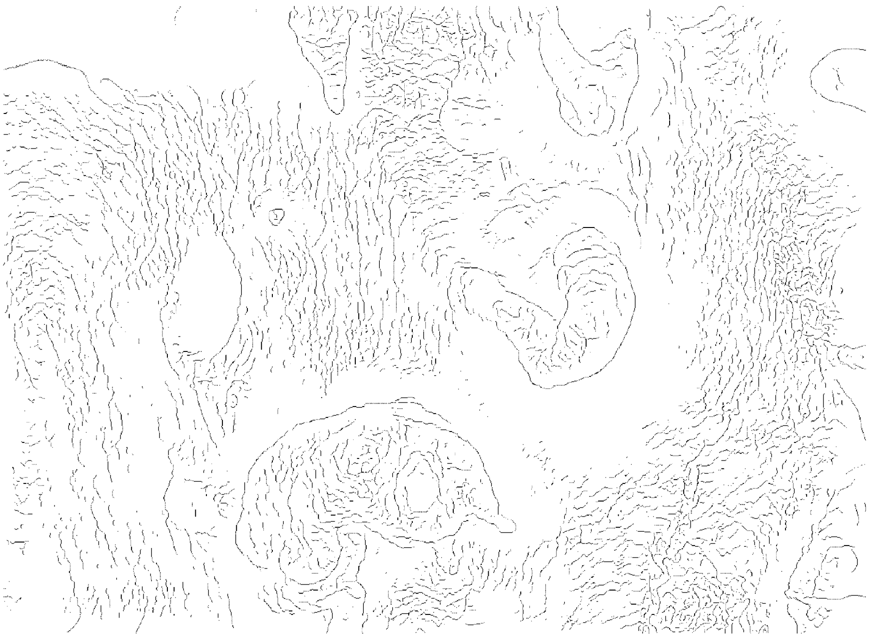

Figure 14. Multiscale edges calculated from the pseudo-gravity transform of Fig. 13 at scale $s = 1$ ($z = z_0$).

the flat background that are not easily seen in the original image, and are quite invisible in the pseudo-gravity transform from which the wavelet transform derives.

Fig. 15 shows the multiscale edges at $s = 5$. We see that much of the fine detail has decayed away, and that only some edges survive. However, it is also apparent that the amplitude of the variation in the original image is not the sole determinant of whether an edge survives or not. Indeed, it is the type of singularity and its depth that determines the decay of the edge as a function of s, and this can lead to weak edges at $s = 1$ that nevertheless persist to larger scales. Although $W_{\max}$ is not shown in Fig. 15, some of these edges are becoming stronger (in the sense that $W_{\max}$ is still increasing with s) at $s = 5$.

Recall that these multiscale edges represent boundaries, or edges in the upward continuation of the field. This allows us to make a connection with 'depth slicing' or separation filtering techniques. The separation filtering view of the upward continuation operator is that it is the product of two successive steps. First, a separation filter is applied to the signal at $s = 0$ to eliminate the spectral content due to shallow sources above

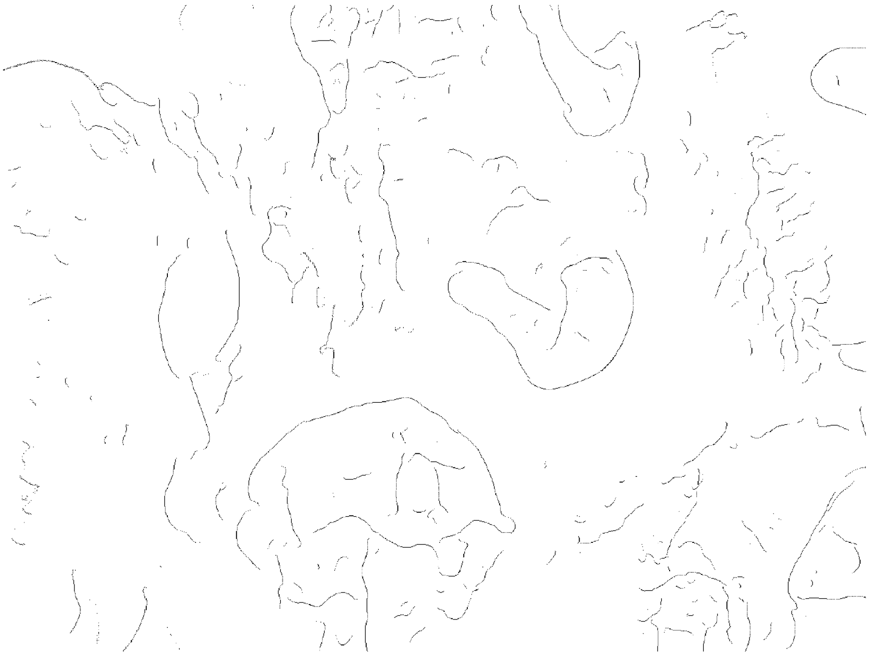

Figure 15. Multiscale edges calculated from the pseudo-gravity transform of Fig. 13 at scale $s = 5$ ($z = 5z_0$).

$z = -s$. It is generally argued that the optimal such filter is the upward continuation operator to height $2s$. This filtered signal is then regarded as an approximation to the field at $s = 0$ arising from sources below $z = -s$. Second, this filtered signal is *downward* continued to depth $z = -s$. The nett filter is upward continuation to height $z = s$, which is then interpreted as an estimation of the field at $z = -s$ after stripping the sources above $z = -s$ (for example Jacobsen 1987). Thus our multiscale edges at a given scale s can be viewed (in the separation filtering interpretation) as the edges present in this notional field at *depth* $z = -s$.

In the separation filtering interpretation, the spectral content of the signal is thought to arise from some form of random-phase, linear superposition of effects arising from an 'ensemble' of possible sources. While this view has merit when considering statistical quantities, such as spectral averages, it is considerably less well justified as a view of all filtering operations of an image. A data image is, after all, but one realization drawn from the set of all possible fields that could arise from the ensemble that is (tacitly) assumed to underlie the analysis. Spectral averaging makes up for this rather dramatic lack of instances (one) by replacing ensemble averages with spectral averages (a kind of ergodicity). Not all manipulations in the Fourier domain admit such an interpretation. In general, it is an error to confuse the spectral properties of a stochastic process with the spectral properties of a single instance of the process unless the spectral property is a statistical quantity being estimated on the basis of an ergodic hypothesis.

Our results for the strength and persistence of edges in the upward continuation are rigorous results pertaining to deterministic features of the source. That is, we interpret the spectral content of the signal as arising from a single instance of a source distribution, which is of course the actual situation.

This difference in worldview leads to different conclusions based upon the same data and processing technique. For example, the rapid decay of an edge as s increases could be due to a deep source with a negative Lipschitz exponent. However, the separation filtering view would regard the decay of the field anomaly associated with the edge as being indicative of a shallow feature. More generally, our analysis demonstrates that the apparent depth of features obtained from the 'depth slicing' method is determined as much by the Lipschitz continuity of the variations in the source as it is by the depth of the source variation.

Maps similar to Figs 14 and 15 have already been produced for exploration purposes in other areas of Australia, and the reaction of experienced interpreters has been positive. If nothing else, these maps represent boundaries of features in the domain of downward-continued separation filterings, and for this reason have a look familiar to interpreters. The next challenge is to present these multiscale edges in a way that illustrates the information about the source structure in an intuitive and comprehensible way.

As a first step towards a multiscale edge inversion tool, 3-D renderings such as Fig. 16 have been produced. Here, the edges produced at different scales are stacked one upon the other, and the amplitude of the wavelet extrema mapped with different colours. This representation of the multiscale edges overlies a 3-D relief rendering of the magnetic map. The vertical variation of the colour of the edges reflects the amplitude variation of the multiscale edge as a function of s, and different styles of colour variation correspond to different types and depths of source variation. The maximum scale edges shown in Figs 16 and 17 are at (approximately) $s = 64$ (800 m). The edges have been restricted to this range to minimize clutter. The edges were actually calculated up to $z = 10$ km (the whole area is approximately 75 km wide). At very large scales, the edges simply separate the high-amplitude regions of the map from low-amplitude regions, and probably do not correspond to any singularity in the source. Research directed towards the steady refinement of such tools is ongoing.

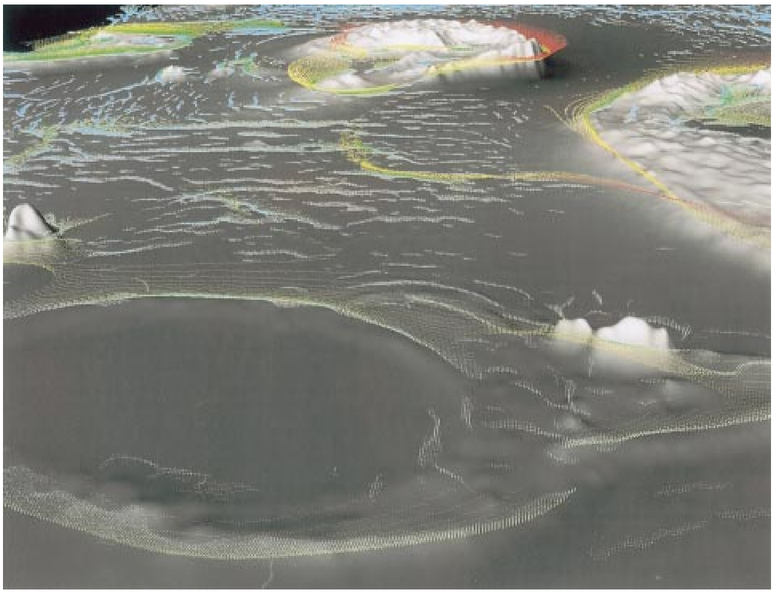

Figure 16. Perspective relief view of the data of Fig. 13 with the multiscale edges suspended in the space above the image. Increasing height above the image represents increased s, in the natural way. The amplitude of the multiscale edge ($W_{\max}$) has been mapped to colour. In this case, blue is low and red is high. The view is from the lower half of the left hand edge of Fig. 13 looking down and slightly to the north of due east.

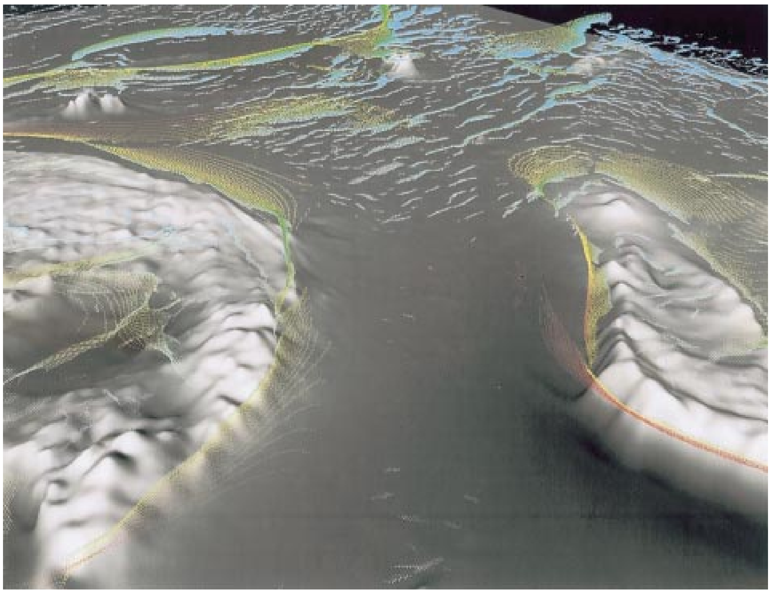

Figure 17. Perspective relief view of the data of Fig. 13 in the same style as Fig. 16. The view is from near the right-hand end of the bottom edge of Fig. 13, looking down and northwest.

7 CONCLUSIONS

Having already summarized most of our results in Sections 3.1.4 and 5, we will devote this section to speculation based upon the foregoing.

First, let us make a clear distinction between the reconstruction problem and the inversion problem. By reconstruction we mean the recovery of $f_0(x, y)$ from a subset of the values of the wavelet transform $W[f_0](s, x, y)$, while the inversion problem is the recovery of $\rho(x, y, z)$.

The recovery of f_0 from the wavelet transform is demonstrated in Appendix B, where it is also shown that a discrete sampling of the scale parameter is also exactly invertible. That the latter is possible demonstrates the redundant information content of the wavelet transform, as is apparent from the fact that the domain of $W[f_0]$ is 3-D, while the domain of f_0 is 2-D. This suggests that knowledge of $W[f_0]$ on some 2-D submanifold of the (s, x, y) domain may be sufficient to reconstruct f_0. A candidate for such a 2-D submanifold is the set of multiscale edges defined by the wavelet transform. Were this to be the case, it would imply that, like the Fourier transform, the multiscale edges re-package the information contained in the image f_0. We have already outlined (in Section 2.1) our reason for suspecting the existence of such an inversion—let us discuss here the consequences in greater detail.

While the Fourier transform illuminates the harmonic content of the function f_0, the multiscale edge representation highlights the positions and types of singularities in the source distribution. Carrying the analogy further, the multiscale edge representation could be used for filtering and processing the image, just as Fourier analysis is currently used. The difference is that the multiscale edge processing would facilitate filtering according to feature type rather than according to frequency content. This possibility opens the door to a range of applications, including data compression, feature extraction, stable downward continuation, layer stripping inversion, and even the improved interpolation of the flight line or traverse data used to construct the data image f_{z_0} in the first place.

Let us hypothesize that we have an image arising from some boundary (call it the main edge), cross-cutting a set of smaller edges or features. Then the multiscale edges can be expected to interact and distort each other. Suppose that the general behaviour of the main edge is clear enough that the corresponding multiscale edge can be isolated and smoothed (even extrapolated to smaller scale!). Then an image could be constructed from the smoothed, interpolated edge in isolation. This image could then be subtracted from the original image, and the resulting smaller features possibly examined free from the effects of the main edge. This is just one example of the many possibilities on offer. The method is, in every respect, as much a signal processing technique as it is a physical theory and basis for inversion. We find this combination quite compelling.

Our preliminary results on the reconstruction issue are showing some promise. The particular wavelets we are using are non-compact, non-orthogonal, non-discrete, and cannot be expressed as a product of functions of x and y. This makes them quite unlike anything of current interest in the wavelet literature. (Indeed, they seem to have been constructed by placing the word 'not' in front of every adjective describing the current focus of study.) However, they are analytic (harmonic) and invariant under Euclidean metric preserving transformations. Analyticity is proving to be advantageous for the reconstruction from subsets of the wavelet domain, while symmetry of the wavelets is leading to particularly neat results regarding the generation of the reconstruction wavelets from a single 'mother' wavelet. We demonstrate the utility of the rotation invariance in Appendix B, where we derive the wavelet inversion transformation from a geophysicist's perspective (from the convolution theorem rather than the frame approach that appeals to mathematicians). This derivation is even easier than for the case of product wavelets, and does not require the signal's analyticity.

The inversion problem (recovery of ρ) is well known to be non-unique, even given complete knowledge of $f_z(x, y)$ for some z in 'free space'. Indeed, it is quite straightforward to characterize *all* the density distributions that could yield a given $f_z(x, y)$. Consider a discrete formulation of the problem. If $f_z(x_i, y_j)$ is given on an $m \times n$ grid, then this leads to mn linear constraints on the Fourier coefficients of the discrete voxel representation $\hat{\rho}(k_{x_i}, k_{y_j}, z_k)$. This is enough information to determine only mn values of $\hat{\rho}(k_{x_i}, k_{y_j}, z_k)$ (the equivalent of one layer of voxels at the same horizontal resolution as the data image). The inversion for p layers of voxels can proceed by searching the set of $(p-1)mn$ solutions in order to find one that matches other constraints (for example, realistic density values or independent drill-hole data) or is in some sense the most likely according to *a priori* information. The problem is that such constraints and likelihood, while often easily expressed in the spatial representation $[\rho(x_i, y_j, z_k)]$, are very awkward to express in the Fourier domain, which is where the search for solutions is best carried out.

The wavelet representation offers a couple of possibilities here. First, the source singularity information gleaned from the multiscale edges (that is, edges in the source) could conceivably be used to set an *a priori* likelihood function on the variation of the source. For example, penalizing variations in the source intensity at points in space not associated with edges would amount to a (soft) constraint tending to make the source variation piecewise-constant. In addition, wavelets are, by nature, quite localized in space. Consequently, spatially local constraints are slightly easier to express in the wavelet domain than in the Fourier domain. This observation is less useful here than it might usually be, since the wavelets we are using have non-compact support. Nevertheless, there may be some merit in expressing both the constraints and the search space in the wavelet basis (a consideration which led us down this path in the first place).

More conventionally, inversion of potential field data seeks an indication of depth, extent and type of source structure based upon the observed field anomalies. Does the current wavelet analysis bear any relation to these techniques, and does it illuminate them in any way? We have mentioned the similar aims and spirit of this analysis, and the Werner- and Euler-based methods. As in the Euler method, the type of singularity located by the wavelet analysis can be an output of the method—in our case, to be deduced directly from the behaviour of $W_{\max}$. In contrast, Werner-based methods treat the type of singularity as an input (for example Hartman *et al.* 1971). We also note in passing that the method outlined by Naudy (1971) is algorithmically quite similar to the current method. In particular, the cross-correlations used by Naudy are closely related to the smoothing function convolution in eq. (2.5). In our present work, we use the Green's function (point-source response), scale the Green's function according to depth, convolve with the image (or profile), differentiate, and adjust the gradient by a factor s to yield the wavelet transform. By contrast, Naudy uses the response of some assumed source geometry (sheet), scales the profile according to depth (by re-sampling), cross-correlates, and does not differentiate. The relationship is clear. The cross-correlations calculated by Naudy are nearly equivalent to our smoothing function convolution $[f * \theta_s]$.

In the case of the Euler method, it is possible to invert also for the 'structural index' (see, for example, Thompson 1982). Indeed, there would seem to be some relationship between the structural index of the source and the Lipschitz regularity of the source. A simple view of the structural index is that it is the exponent governing the rate of decay with distance from the source. In the case of gravitational acceleration, the acceleration decays as the inverse square of the distance to a point source. Hence a point source has a structural index of two in this case. It is easy to see that the structural index is influenced by two factors. First, there is the rate of decay of the Green's function with distance. Second, there is the rate of increase of source material within the field of influence of a Green's function that broadens with increasing distance to source. So, for a line source, the gravitational acceleration Green's function decays as the inverse square of distance, but the amount of source material to be found within the field of influence of the Green's function is increasing as the first power of the distance. The nett result is an acceleration that decays as the inverse first power of distance, and hence a structural index of one for a line source. The numerical similarity to the Lipschitz regularity is apparent.

The forgoing indicates that the structural index corresponds to the exponent governing the rate of decay of our smoothing function convolution, $[f * \theta_s]$, as a function of s. Now note that, in eq. (2.2), we differentiate the smoothing function to form the basic wavelet, and then scale the wavelet in (2.4) to form the wavelet family. The order of the differentiation and scaling is important. The effect of doing the derivative first is the extra factor of s that multiplies the gradient in the formula (2.5) for the wavelet transform. The upshot of this is that the wavelet transform amplitudes always scale the same way as the smoothing function convolution. Indeed, if we were to go to second-order wavelets, we would differentiate the smoothing function twice, and then scale the resulting functions. This means that our wavelet transforms always scale in a manner determined by the source and the smoothing function, independent of the order of the transform.

Let us contrast this with the structural index. When dealing with dipole sources, one observes an inverse cubic decay rate for a 'point dipole'. Thus, the structural index of the dipole point source is taken as three. This is more a consequence of the fact that the Green's function of the dipole field is the gradient of the Green's function of the monopole field than it is a characteristic of the source (although we concede that this point is moot). In fact, the structural index in this case has been obtained by observing the decay rate of a differentiation of the monopole field. This is what we would obtain from the wavelet-style analysis were we to first scale the smoothing function, and then differentiate to obtain the wavelet basis. In effect, we would lose the factor of s multiplying the gradient, and the relationship between the source Lipschitz regularity and the wavelet decay rate would be disturbed.

The structural index is also related to the degree of homogeneity of the observed signal or field measurement. In this connection, the relationship between the behaviour of the Poisson-based wavelet transform, and the structural index has recently been pointed out by Moreau *et al.* (1997).

We see that re-casting potential field theory into wavelet terms yields novel signal processing methods, illuminates traditional deconvolution methods, and offers some possibilities for further progress in the area of full inversion and data fusion. Whatever the future of the full inversion approach, and the impact of wavelet analysis upon it, it would seem that the characterization of singular source structures and the possibilities offered by a solution to the reconstruction problem will ensure that wavelet analysis, as a signal processing method, has a sound future in the processing of potential field data. We hope to have outlined the basics of the application of continuous, non-compact, non-orthogonal, non-product wavelets to the processing of potential field data, and have risked some rather speculative conjectures in order to alert the reader to some of the possibilities. Again, we emphasize that we do not claim any superior performance or new information extraction properties for these methods. Rather, we seek to point out that these methods shed light upon many of the existing techniques, by casting those methods in a more general framework. In doing so, the method shows a way forward, not only for the wavelet analysis of potential field data, but also for those traditional methods that are in some way equivalent or closely related.

ACKNOWLEDGMENTS

We wish to acknowledge Fractal Graphics Pty Ltd for financial support, and S. Mallat for permission to use the software that prompted us to look more deeply into these matters. We also thank the Geological Survey of Victoria for access to the data behind Fig. 13. Work reported here was conducted as part of the Australian Geodynamic Cooperative Research Centre and this paper is published with the permission of the Director, AGCRC.

REFERENCES

Al-Chalabi, M., 1971. Some studies relating to nonuniqueness in gravity and magnetic inverse problems, *Geophysics,* **36,** 835–855.

Blakely, R.J., 1995. *Potential Theory in Gravity and Magnetic Applications,* Cambridge University Press, Cambridge.

Blakely, R.J. & Simpson, R.W., 1986. Approximating edges of source bodies from magnetic or gravity anomalies, *Geophysics,* **51,** 1494–1498.

Boschetti, F., Horowitz, F.G. & Hornby, P., 1998. Ambiguity analysis and the constrained inversion of potential field data, *Geophys. J. Int.,* submitted.

Boschetti, F., Dentith, M.C. & List, R.D., 1996. Inversion of seismic refraction data using Genetic Algorithms, *Geophysics,* **61,** 1715–1727.

Cordell, L., 1985. A stripping filter for potential field data, in *SEG, 55th Ann. Mtg, Abstracts with Programs,* 217–218.

Federer, H., 1969. *Geometric Measure Theory,* Springer-Verlag, Berlin.

Grauch, V.J. & Cordell, L., 1987. Limitations of determining density or magnetics boundaries from the horizontal gradient of gravity or pseudogravity data, *Geophysics,* **52,** 118–121

Hartman, R.D., Teskey, D.J. & Friedberg J.L., 1971. A system for rapid aeromagnetic interpretation, *Geophysics,* **36,** 891–918.

Holschneider, M. & Tchamitchian, P., 1989. Regularite locale de la function 'non-differentiable' de Riemann, in *Les Ondelettes en 1989,* pp. 102–123, ed. P.G. Lemarie, Lecture notes in Mathematics, Springer-Verlag.

Hummel, R. & Moniot, R., 1989. Reconstructions from zero crossings in scale space, *IEEE Trans. on Acoustics, Speech and Signal Processing,* **37,** 2111–2130.

Hwang, W.-L. & Mallat, S., 1993. Characterisation of self-similar multifractals with wavelet maxima, *Technical report 641,* Courant Inst. of Math. Sci., Comp. Sci. Department, New York University, New York.

Jacobsen, B.H., 1987. A case for upward continuation as a standard separation filter for potential field maps, *Geophysics,* **52,** 1138–1148.

Jaffard, S., 1997. Multifractal formalism for functions part 1: results valid for all functions. *SIAM J. Math. Anal.,* **28,** 944–970.

Kaiser, G., 1994. *A Friendly Guide to Wavelets,* Birkhauser, Boston.

Mallat, S., 1991. Zero-crossing of a wavelet transform, *IEEE Trans. Information Theory,* **37,** 1019–1033.

Mallat, S. & Zhong, S., 1992. Characterisation of signals from multiscale edges, *IEEE Trans. Pattern Analysis and Machine Intelligence,* **14,** 710–32.

Mareschal, J.C., 1985. Inversion of potential field data in Fourier transform domain, *Geophysics,* **50,** 685–691.

Mikhailov, V.P., 1978. *Partial Differential Equations,* MIR Publishers, Moscow.

Moreau, F., Gibert, D., Holschnider, M. & Saracco, G., 1997. Wavelet analysis of potential fields, *Inverse Problems,* **13,** 165–178.

Naudy, H., 1971. Automatic determination of depth on aeromagnetic profiles, *Geophysics,* **36,** 717–722.

Piech, M.A. & Piech, K.R., 1990. Fingerprints and fractal terrain, *Math. Geol.,* **22,** 457–485.

Rogers, C.A., 1970. *Hausdorff Measure,* Intro., 1.1, 1.2, 3.2, 5.1, 5.2, 8.5, Cambridge University Press, Cambridge.

Telford, W.M., Geldart, L.P., Sheriff, R.E. & Keys, D.A., 1976. *Applied Geophysics,* Cambridge University Press, Cambridge.

Thompson, D.T., 1982. EULDPH: A new technique for making computer-assisted depth estimates from magnetic data, *Geophysics,* **47,** 31–37.

Werner, S., 1953. Interpretation of magnetic anomalies at sheet-like bodies, *Sveriges Geologiska Undersok, Arsbak,* **43,** no. 6, series C, no. 508.

Yuille, A.L. & Poggio, T.A., 1986. Scaling theorems for zero crossing, *IEEE Trans. Pattern Analysis and Machine Intelligence,* **8,** 15–25.

APPENDIX A: WAVELET SCALING AND LIPSCHITZ COEFFICIENTS

We consider functions satisfying eq. (2.6a) with $\alpha \in (0, 1]$, in particular the continuity and differentiability of such functions. From (2.6a) we see that, given an $\varepsilon > 0$, we can find a $\delta = (\varepsilon/C)^p > 0$ (where $p = \alpha^{-1}$) such that $|x_0 - x_1| < \delta$ implies that $|f(x_0) - f(x_1)| \leq \varepsilon$. Thus f is uniformly continuous on (a, b) and can be extended uniquely to a continuous function on $[a, b]$. This deals with the question of continuity, and from this continuity we can conclude several more facts. For example, since f is a continuous map from a compact, connected closed topological space into a metric space, its image is closed, compact, and totally bounded. Consequently, the image of $[a, b]$ under f has finite Lesbegue measure.

Let us now turn to the differentiability of a function F for which $\mathrm{LUR}(F) \in (0, 1]$. (The intent is that F is the primitive function mentioned in Section 2.3 in connection with the definition of negative Lipschitz exponents.) Let $I \equiv [x_0, x_r]$ be an interval $I \subset (a, b)$ and $\{x_i\}_{i=0}^{r}$ a set of points in I such that $x_{i-1} < x_i$, and consider the quantity (called the total variation of F on I)

$$V_I(F) = \sup_{\{x_i\}} \sum_i |F(x_i) - F(x_{i-1})|, \tag{A1}$$

where the supremum is over all such subdivisions of I. If $N(F, I, y)$ is the (possibly infinite) number of values x in an interval $I \subset (a, b)$ such that $F(x) = y$, then

$$V_I(F) = \int_{f(I)} N(F, I, y)\nu^1(dy), \tag{A2}$$

whenever F is continuous. (v^1 is the usual Lesbegue measure in one dimension; for example Federer 1969, p. 177.) However, we have already established that if $\mathrm{LUR}(F) \in (0, 1]$ then F is continuous, can be extended to the closed interval, and that its image has finite Lesbegue measure. Suppose that $N(F, [a, b], y)$ is finite for almost all y. If this is the case then $V_{[a,b]}(F)$ is finite and hence F is of bounded variation on $[a, b]$. Summarizing, if $\mathrm{LUR}(F) \in (0, 1]$ then F is continuous on the closed interval $[a, b]$ and moreover, if it does not oscillate too much, it is also of bounded variation.

The set of linear functionals on the real line is isomorphic to the real line, and hence the real numbers are reflexive as a Banach space. Under these conditions (bounded variation over a reflexive Banach space), there exist Borel regular measures μ_F^+ and μ_F^- such that

$$F(x) - F(a) = \mu_F^+([a, x]) - \mu_F^-([a, x]) = \int_{[a,x]} dF , \qquad \text{(A3)}$$

where the integral is understood in the Riemann–Stieltjes sense. Stated in physical terms, a function F such that $\mathrm{LUR}(F) \in (0, 1]$ and of bounded variation induces a source distribution that can be denoted by dF, and the associated field can be understood in terms of the Riemann–Stieltjes integrals of the Green's function with respect to dF.

The $\mu_F^\pm$ are absolutely continuous with respect to Lesbegue measure if and only if F is absolutely continuous (Federer 1969, pp. 164–166). In the absolutely continuous case we can define the density

$$f = F_{,1} = \frac{d\mu_f^+}{dv^1} - \frac{d\mu_f^-}{dv^1} , \qquad \text{(A4)}$$

where the derivatives are in the Radon–Nikodym sense. This density Lesbegue integrates back to F, making it a valid candidate for the derivative of F. Mixing various notations with which the reader may be more familiar, we have

$$\mu_F^+(dx) - \mu_F^-(dx) = F_{,1} v^1(dx) = D_x F dx = f dx = dF . \qquad \text{(A5)}$$

One can view the Lipschitz condition as a condition on the first derivative of F (when it exists). Substituting $|F(x_1) - F(x_0)| = |f(\xi)| \|x_1 - x_0|$ (for some $\xi \in [x_0, x_1]$) in the Lipschitz condition yields a bound of $\alpha - 1$ on the exponent governing any divergence of the first derivative f of F (when it exists). This same number $(\alpha - 1)$ is taken as the Lipschitz exponent of f, as outlined in Section 2.3. In the case when the first derivative does not exist in the conventional sense, we associate the exponent $\alpha - 1$ with dF, and the measures $\mu_F^\pm$. That is, $\mathrm{LUR}(dF) = \mathrm{LUR}(\mu_F^\pm) = \mathrm{LUR}(F) - 1$. Even if F is not absolutely continuous, (2.6a) can still be used to bound $\mu_F^\pm$ with a Hausdorff measure of dimension α, and so dF can be expressed as some superposition of fractional dimensional measures of dimension bounded below by α. Indeed, the right-hand side of (2.6a) can be replaced with $Ch(|x_0 - x_1|)$ for some suitable positive function h vanishing at zero, increasing and right continuous, the functions $h(\ell) = \ell^\alpha$ being a special case (Rogers 1970). In this more general setting, the 'dimensionality' of the source is parametrized by functions h (or, more correctly, by equivalence classes of functions) rather than by a single number.

The above indicates that the wavelet methods can be used to study much less regular sources and field configurations, multifractal sources being a special case (Hwang & Mallat 1993). Moreover, it is possible to define 'Lipschitz exponents' greater than unity in terms of polynomial approximations, and to generalize the present results further (Jaffard 1997).

The above discussion was for functions and measures of a single variable. However, we have kept the language general enough to make the passage to higher dimensions reasonably clear, provided one does not attempt to be too general. In a bounded singly connected open domain of $\mathbf{R}^n$ we require

$$|f(\mathbf{x}_0) - f(\mathbf{x}_1)| \le C\|\mathbf{x}_0 - \mathbf{x}_1\|^\alpha , \qquad \text{(A6)}$$

where $\|\mathbf{x}\|$ is the usual Euclidean norm in $\mathbf{R}^n$. If $\alpha \in (0, 1]$ then f is again uniformly continuous, and extends to a continuous function on the closure of the domain. It is also worth noting that the previous discussion applies to each line segment in the domain.

With the above as background we make a slight modification to the forward inference in Theorem 1 for the case of a logarithmic singularity in the first derivative of the function being transformed. We also consider more than one dimension.

Theorem 2. Let a real-valued function f satisfy the condition

$$|f(\mathbf{x}) - f(\mathbf{y})| \le C\|\mathbf{x} - \mathbf{y}\|^\alpha |\log\|\mathbf{x} - \mathbf{y}\||^\beta , \qquad \text{(A7)}$$

where $\alpha \in (0, 1]$ and $\beta \in \{0, 1\}$ then

$$M[f](s, \mathbf{x}) \le \begin{cases} K_1 s^\alpha & \beta = 0 \\ K_1 s^\alpha |\log(s)| + K_2 s^\alpha & \beta = 1 \end{cases} . \qquad \text{(A8)}$$

The demonstration of Theorem 2 follows closely that of Theorem 1 (Holschneider & Tchamitchian 1989).

Consider the wavelet $\boldsymbol{\psi}$ such that $\int \boldsymbol{\psi}\, d\mathbf{x} = 0$. Then

$$\begin{aligned} \mathbf{W}[f](s, \mathbf{b}) &= s^{-1} \int_{\mathbf{R}^n} \boldsymbol{\psi}[(\mathbf{x} - \mathbf{b})/s] f(\mathbf{x})\, d\mathbf{x} \\ &= s^{-1} \int_{\mathbf{R}^n} \boldsymbol{\psi}[(\mathbf{x} - \mathbf{b})/s][f(\mathbf{x}) - f(\mathbf{b})]\, d\mathbf{x} . \end{aligned} \qquad \text{(A9)}$$

Taking the norm of both sides, bounding the right and inserting (A7), and then making the substitution

$$\mathbf{u}(\mathbf{x}) = s^{-1}(\mathbf{x} - \mathbf{b}) , \qquad \text{(A10)}$$

one finds

$$\begin{aligned} M[f](s, \mathbf{b}) &= \left\| s^{-n} \int_{-\infty}^{\infty} [f(\mathbf{x}) - f(\mathbf{b})] \boldsymbol{\psi}[(\mathbf{x} - \mathbf{b})/s]\, d\mathbf{x} \right\| \\ &\le s^{-n} \int_{-\infty}^{\infty} \|(\boldsymbol{\psi} \circ \mathbf{u})(\mathbf{x})\|\, |f(\mathbf{x}) - f(\mathbf{b})|\, d\mathbf{x} \\ &\le C \int_{-\infty}^{\infty} \|(\boldsymbol{\psi} \circ \mathbf{u})(\mathbf{x})\|\, \|\mathbf{x} - \mathbf{b}\|^a |\log\|\mathbf{x} - \mathbf{b}\||^\beta s^{-n}\, d\mathbf{x} \\ &= C \int_{-\infty}^{\infty} \|\boldsymbol{\psi}(\mathbf{u})\|\, \|s\mathbf{u}\|^\alpha |\log\|s\mathbf{u}\||^\beta\, d\mathbf{u} \\ &\le C s^\alpha \int_{-\infty}^{\infty} \|\mathbf{u}\|^\alpha \|\boldsymbol{\psi}(\mathbf{u})\| (|\log(s)| + |\log\|\mathbf{u}\||)^\beta\, d\mathbf{u} \\ &= \begin{cases} K_1 s^\alpha & \beta = 0 \\ K_1 s^\alpha |\log(s)| + K_2 s^\alpha & \beta = 1 . \end{cases} \end{aligned} \qquad \text{(A11)}$$

Let us now consider densities with negative exponents, first in the 1-D case. Suppose $\mathrm{LUR}(F) = (\alpha + 1) \in (0, 1]$ and that F has bounded variation. Then integrating by parts and applying a 1-D version of the substitution (A10) to the Riemann–Stieltjes integral (noting also that $\int \psi_{,1}\, dx = 0$, we find

$$\begin{aligned}
|W[dF](s, b)| &= s^{-1} \left| \int_{-\infty}^{\infty} \psi \circ u \, dF \right| \\
&\leq s^{-1} \left| \int_{-\infty}^{\infty} F \, d(\psi \circ u) \right| \\
&= s^{-1} \left| \int_{-\infty}^{\infty} F \circ u^{-1} \, d\psi \right| \\
&= s^{-1} \left| \int_{-\infty}^{\infty} F(su + b) \psi_{,1}(u) \, du \right| \\
&= s^{-1} \left| \int_{-\infty}^{\infty} [F(su + b) - F(b)] \psi_{,1}(u) \, du \right| \\
&\leq s^{-1} \int_{-\infty}^{\infty} |F(su + b) - F(b)| |\psi_{,1}(u)| \, du \\
&\leqslant C s^{\alpha} \int_{-\infty}^{\infty} |\psi_{,1}(u)| \, |u|^{\alpha + 1} \, du \\
&= K s^{\alpha} .
\end{aligned} \tag{A12}$$

This is consistent with the definition $\mathrm{LUR}(dF) = \mathrm{LUR}(F) - 1$. Our final task is to give an indication of how this last manipulation can be generalized to n dimensions. Rather than attempt anything fancy, we will make a simple and direct assumption about the distribution of source material that allows immediate application of (A12).

Suppose that μ is some Borel regular measure, and define the function F by

$$F(\mathbf{b}, r) = \mu(B(\mathbf{b}, r)) = \int_{E \cap B(\mathbf{b}, r)} \mu(d\mathbf{x}) , \tag{A13}$$

where E is the support of the measure and $B(\mathbf{b}, r)$ denotes a closed ball centred at $\mathbf{b}$ of radius r. Such an F is clearly a non-decreasing function of r, and can be used to define an integral on the positive real line, which we will denote by $\int_0^{\infty} dF_{\mathbf{b}}$. This integral is the one induced on the radial coordinate under transformation of the μ integral to polar coordinates about $\mathbf{b}$. Let us suppose that

$$F(\mathbf{b}, r) \leq C r^{n + \alpha} \tag{A14}$$

in some connected domain of $\mathbf{R}^n$. Then we can consider the Lipschitz uniform regularity of μ in a domain as the maximum α satisfying (A14) for all $B(\mathbf{b}, r)$ in that domain. In this case we define $\mathrm{LUR}(\mu) = \alpha$. The case $\mathrm{LUR}(\mu) = -n$ is intended to assert that μ has atomic measure contributions.

Noting that $\|\boldsymbol{\psi}(\mathbf{x})\|$ is invariant under rotations (by virtue of the symmetry of θ) we can define

$$m(r) = \| \boldsymbol{\psi}(r\mathbf{n}) \| , \tag{A15}$$

which is the same function of r, independent of the unit vector $\mathbf{n}$. We also note that $\lim_{r \to 0} m(r) = 0$. By first transforming to polar coordinates about $\mathbf{b}$, then integrating by parts we find that

$$\begin{aligned}
M[\mu](s, \mathbf{b}) &= \left\| s^{-n} \int_E \boldsymbol{\psi} \circ \mathbf{u} \mu(d\mathbf{x}) \right\| \\
&\leq s^{-n} \int_E \| \boldsymbol{\psi} \circ \mathbf{u} \| \mu(d\mathbf{x}) \\
&= s^{-n} \int_0^{\infty} m(r/s) \, dF_{\mathbf{b}} \\
&\leq s^{-n} \left| \int_0^{\infty} F(\mathbf{b}, sv) m_{,1}(v) \, dv \right| \\
&\leq C s^{\alpha} \int_0^{\infty} v^{n+\alpha} |m_{,1}(v)| \, dv \\
&= K s^{\alpha} ,
\end{aligned} \tag{A16}$$

provided the boundary terms in the integration by parts vanish, and that the various integrals converge.

It remains to clarify the relationship between these results and eq. (2.23). First, let us consider the vertical acceleration $f_z^{\delta}(x, y)$ at height z generated by a planar source $\rho(x, y, z)\delta(z - z')$ at depth $z' < 0$. Let us consider a wavelet transform of the planar source based upon scale zero coinciding with $z = z'$ rather than $z = 0$. Then, if the planar source has Lipschitz uniform regularity α it follows from the forgoing that

$$\frac{z - z'}{z_0} \| \nabla f_z^{\delta} \| \leq K \left(\frac{z - z'}{z_0} \right)^{\alpha} . \tag{A17}$$

Introducing $s = z/z_0$ and $s' = -z'/z_0$ as in (2.23), and after some minor rearrangement, (A17) becomes

$$\| \nabla f_z^{\delta} \| \leq K (s + s')^{\alpha - 1} . \tag{A18}$$

Moving zero scale back to $z = 0$, and considering the $z = 0$ based wavelet transform of the acceleration f_0^{δ} at zero height, we find from the fundamental relations (2.22a,b) that

$$\begin{aligned}
M[f_0^{\delta}](s, \mathbf{x}) &= (z/z_0) \| \nabla f_z^{\delta} \| \\
&\leq K (s + s')^{\alpha - 1} s ,
\end{aligned} \tag{A19}$$

which is just (2.23). In essence, we have used the wavelet scaling result to infer the way that $\| \nabla f_z^{\delta} \|$ scales with distance from the *causative singularity*, and then used the relationship between the $z = 0$ based wavelet transform and the gradient to ascertain the scaling behaviour of the $z = 0$ based wavelet transform of f_0^{δ} (in the absence of other sources of course). Note also the underlying consistency here. Even if the planar source has negative Lipschitz regularity, the field f_0^{δ} at $z = 0$ is smooth. This is consistent with the s dependence of (A19), which predicts vanishing wavelet amplitudes as s goes to zero. However, the mark of the singularity at $z = z'$ underlying the smooth field at $z = 0$ is still betrayed in the behaviour of (A19), and might also be coaxed out of hiding by taking a different definition of zero scale. Such an endeavour would first require that all other singularities in the field that lie above z' be removed before attempting to calculate the new wavelet transform based upon zero scale at z'.

This leads to our current research, which seeks to identify zero and negative Lipschitz regular sources at $z = 0$, isolates these features, reconstructs the signal corresponding to just these features, and subtracts these sources' effects from the original data. The residual data, with the shallow sources stripped, might then be suitable for the construction of wavelet transforms whose zero scale is below the $z=0$ plane. This process of stripping away sources can in principle be repeated. The idea is not as difficult to implement as it first appears, since the multiscale edges undergo a characteristic 'trifurcation' at the depth corresponding to the source position. This characteristic topological behaviour, together with the tell-tale scale behaviour of the multiscale edges, makes simple singularities rather easy to pick.

APPENDIX B: INVERSE WAVELET TRANSFORM

In this appendix, we present a very simple derivation of the inverse wavelet transform. We have taken the derivation from Kaiser (1994), re-cast it in the language of convolutions, and made some simplifications that arise in the case of Euclidean symmetry of the smoothing function.

Taking Fourier transforms of the basic wavelet transformation equation, we find that

$$\hat{\mathbf{W}}[f_0](s,\mathbf{k})=\hat{\boldsymbol{\psi}}_s(\mathbf{k})\hat{f}_0(\mathbf{k})\,. \tag{B1}$$

One begins by removing the vector nature of (B1). Taking inner products, one finds

$$(\hat{\boldsymbol{\psi}}_s(\mathbf{k}),\hat{\mathbf{W}}[f_0](s,\mathbf{k}))_{C^2}=\|\hat{\boldsymbol{\psi}}_s(\mathbf{k})\|^2_{C^2}\hat{f}_0(\mathbf{k})\,, \tag{B2}$$

where the inner product and norm are the usual Euclidean norm of 2-D linear vector spaces over the complex numbers. Clearly, we could divide (B2) by the factor $\|\hat{\boldsymbol{\psi}}_s(\mathbf{k})\|^2_{C^2}$, and play all the usual tricks (which hardly ever work very well) to deal with possible division by small numbers. A neater trick is to integrate both sides over s, thereby dramatically reducing the chances of dividing by zero. That is,

$$\int_0^\infty(\hat{\boldsymbol{\psi}}_s(\mathbf{k}),\hat{\mathbf{W}}[f_0](s,\mathbf{k}))_{C^2}\frac{ds}{s}=\int_0^\infty\|\hat{\boldsymbol{\psi}}_s(\mathbf{k})\|^2_{C^2}\hat{f}_0(\mathbf{k})\frac{ds}{s}\,. \tag{B3}$$

Inversion practitioners might like to add this to their bag of tricks. The simplicity is a little misleading, as there are a few steps to be negotiated yet, and we have already engaged in a little 'rabbit pulling' in our choice of the measure ds/s. The choice relates to invariant measures on a product of the symmetry group of the wavelet and the scaling group (for example Kaiser 1994; Federer 1969). Continuing,

$$\hat{f}_0(\mathbf{k})=Y(\mathbf{k})^{-1}\int_0^\infty(\hat{\boldsymbol{\psi}}_s(\mathbf{k}),\hat{\mathbf{W}}[f_0](s,\mathbf{k}))_{C^2}\frac{ds}{s}\,, \tag{B4}$$

where

$$\begin{aligned}Y(\mathbf{k})&=\int_0^\infty\|\hat{\boldsymbol{\psi}}_s(\mathbf{k})\|^2_{C^2}\frac{ds}{s}\\&=\int_0^\infty 4\pi^2 s\|\mathbf{k}\|^2\exp(-4\pi\|\mathbf{k}\|s)\,ds\\&=\frac{1}{4}\,.\end{aligned} \tag{B5}$$

It is in arranging for the identity (B5) that one makes the various choices mentioned earlier. Note in particular the pivotal importance of the rotational symmetry of the integrand, the scaling property $\|\hat{\boldsymbol{\psi}}_s(r\mathbf{k})\|^2_{C^2}=\|\hat{\boldsymbol{\psi}}_{rs}(\mathbf{k})\|^2_{C^2}$ and the way the measure ds/s does nothing to disturb the integral under a change in scale. These factors work together to make the integration possible, and, moreover, lead to a scale-invariant Y. See, in particular, Kaiser (1994) for less stringent conditions on Y that nevertheless lead to 'nice' inversion formulae, and conditions on the basis, and s-measure that lead to these inversions.

Performing the inverse Fourier transform yields

$$f_0(\mathbf{x})=4\int_0^\infty\frac{ds}{s}\int_{\mathbf{R}^2}(\boldsymbol{\psi}_s(\mathbf{u}-\mathbf{x}),\mathbf{W}[f_0](s,\mathbf{u}))_{\mathscr{R}^2}\,d\mathbf{u}\,. \tag{B6}$$

The reconstructing wavelet [$\boldsymbol{\psi}^{s,\mathbf{u}}(\mathbf{x})$ in Kaiser's notation] associated with scale s and point $\mathbf{u}$ [$\boldsymbol{\psi}^{s,\mathbf{u}}(\mathbf{x})=4\boldsymbol{\psi}_s(\mathbf{u}-\mathbf{x})$ in (B6)] is, up to a constant factor, just the original analysing wavelet [$\boldsymbol{\psi}_{s,\mathbf{u}}(\mathbf{x})=\boldsymbol{\psi}_s(\mathbf{u}-\mathbf{x})$]. More generally, the reconstructing wavelets are given by

$$\hat{\boldsymbol{\psi}}^{s,\mathbf{u}}(\mathbf{k})=Y(\mathbf{k})^{-1}\hat{\boldsymbol{\psi}}^*_s(\mathbf{k})\exp(-2\pi i\mathbf{k}\cdot\mathbf{u})\,, \tag{B7}$$

which are not always expressible as translations and scalings of a single function when Y is not constant. (Note that $\boldsymbol{\psi}^*$ denotes the complex conjugate of $\boldsymbol{\psi}$.)

One can go on to calculate the reproducing kernel for the wavelet transform function space, whose defining property is

$$\hat{\mathbf{W}}[f_0](s',\mathbf{k})=\int_0^\infty\hat{\mathbf{K}}(s',s,\mathbf{k})\hat{\mathbf{W}}[f_0](s,\mathbf{k})\frac{ds}{s}\,. \tag{B8}$$

The matrix-valued kernel has matrix elements

$$[\hat{\mathbf{K}}(s',s,\mathbf{k})]^\alpha_\beta=\psi^\alpha_{s'}(\mathbf{k})Y(\mathbf{k})^{-1}\hat{\psi}^{\gamma*}_s(\mathbf{k})\mathsf{g}_{\gamma\beta}\,, \tag{B9}$$

where the metric tensor g is Euclidean in the present case. In the spatial domain with constant Y, one obtains

$$K^\alpha_\beta(s',\mathbf{x}'|s,\mathbf{x})=4\int_{\mathbf{R}^2}\psi^\alpha_{s'}(\mathbf{x}'-\mathbf{u})\psi^\beta_s(\mathbf{x}-\mathbf{u})\,d\mathbf{u}\,. \tag{B10}$$

One of the most useful properties of the reproducing kernel is that it is an orthogonal projection into the space of wavelet transforms, and can therefore be used to find a function whose wavelet transform is the nearest (in the L^2 sense) to some desired form (for example Kaiser 1994).

Finally, consider the case when the wavelet transform is known at a discrete set of scales

$$s_m=\sigma^m\,,\quad \sigma>1\,. \tag{B11}$$

Then for some function g which is integrable in the following sense, we have

$$\begin{aligned}\int_0^\infty g(s)\frac{ds}{s}&=\int_0^\infty g(s)\,d(\log s)\\&\approx\log(\sigma)\sum_{m=-\infty}^{\infty}g(\sigma^m)\,,\end{aligned} \tag{B12}$$

where the summation approximation is just the Riemann sum corresponding to the integral. Repeating the previous derivation, but summing (B2) over discrete scales in the manner of (B12) instead of integrating, one finds that

$$\hat{f}_0(\mathbf{k}) = \log(\sigma) Y_\sigma(\mathbf{k})^{-1} \sum_m (\hat{\boldsymbol{\psi}}_{\sigma^m}(\mathbf{k}), \hat{\mathbf{W}}[f_0](\sigma^m, \mathbf{k}))_{C^2}\,. \tag{B13}$$

Note that this is *exact*. The fact that the summation is an approximation to the integral is irrelevant to the fact that (B13) follows as an identity from (B2) after summing both sides. The function space basis metric is now

$$Y_\sigma(\mathbf{k}) = \log(\sigma) \sum_m \|\hat{\boldsymbol{\psi}}_{\sigma^m}(\mathbf{k})\|^2_{C^2}\,, \tag{B14}$$

which is a self-similar, scale periodic function. The first step is to apply a logarithmic transformation to extract a normal periodic function from the scale periodic function. The periodic function then possesses a Fourier series expansion. Some welcome cancellations occur in the evaluation of the series coefficients, and the final result requires the evaluation of an elementary looking, but very stubborn integral. The final result is

$$Y_\sigma(\mathbf{k}) = Q_{\log(\sigma)}\left(\frac{\log(4\pi\|\mathbf{k}\|)}{\log(\sigma)}\right), \tag{B15}$$

where the corresponding periodic function's Fourier series is

$$Q_\gamma(b) = \frac{1}{4} \sum_{j=-\infty}^{\infty} \Gamma\left(2 - \frac{2\pi j}{\gamma} i\right) \exp(2\pi i j b) \tag{B16}$$

and Γ is the gamma function. Writing the gamma function in terms of elementary functions we get

$$Q_\gamma(b) = \frac{1}{4} + \frac{1}{2} \sum_{j=1}^{\infty} \left(\frac{\frac{2\pi^2}{\gamma} j \left(1 + \frac{4\pi^2}{\gamma^2} j^2\right)}{\sinh\left(\frac{2\pi^2}{\gamma} j\right)} \right)^{1/2} \cos[2\pi j b + \mathrm{Arg}(\Gamma_j)]\,. \tag{B17}$$

For $\sigma = \mathrm{e}$, the first non-constant term is already $O(10^{-3})$, and the terms decrease very rapidly with j. They also vanish rapidly as γ tends to zero (i.e. σ approaching 1). For our purposes, where geophysical processes such as upward continuation are seldom within 5 per cent of the actual value, $Y_\sigma(\mathbf{k})$ is as good as constant for any reasonable choice of σ. If we approximate $Y_\sigma(\mathbf{k})$ with its continuum value of one-quarter, we again obtain a set of inverse wavelets from scalings and translations of a single pair of wavelets. We conclude by simply quoting the results for the inverse transformation and the reproducing kernel in the case of discrete scale:

$$\hat{f}_0(\mathbf{k}) = -\frac{2\pi i \log(\sigma)}{Y_\sigma(\mathbf{k})} \sum_{m=-\infty}^{\infty} \sigma^m \times \exp[-2\pi\sigma^m \|\mathbf{k}\|](\mathbf{k}, \hat{\mathbf{W}}[f_0](\sigma^m, \mathbf{k}))_{C^2} \tag{B18}$$

and

$$[\hat{\mathbf{K}}(m', m, \mathbf{k})]^\alpha_\beta = \frac{4\pi^2 \log(\sigma)}{Y_\sigma(\mathbf{k})} k_\alpha k_\beta \sigma^{(m+m')} \times \exp[-2\pi(\sigma^m + \sigma^{m'})\|\mathbf{k}\|]\,. \tag{B19}$$