Approximating scattered data with discontinuities *

Erlend Arge and Michael Floater

February 1994

Abstract. A method is presented for approximating scattered data by a function defined on a regular two-dimensional grid. It is required that the approximation is discontinuous across given curves in the parameter domain known as faults. The method has three phases: regularisation, local approximation and extrapolation. The main emphasis is put on the extrapolation which is based on a matrix equation which minimises second order differences. By approximating each fault by a set of line segments parallel with one of the axes, it is simple to introduce natural boundary conditions across the faults. The resulting approximation has, as expected, discontinuities across faults and is smooth elsewhere. The method is stable even for large data sets.

Keywords. scattered data approximation, regular grids, smoothing techniques, faults

§1. Introduction

In a previous paper [1], a three-stage method for approximating scattered data by grid functions was described, where a *grid function* is a function defined on the nodes of some regular two-dimensional grid Ω embedded in $\mathbb{R}^2$.

Grid functions are common means for representing surfaces in many branches of the technical sciences. Examples are terrain modelling in the mapping industries and the modelling of geological surfaces in oil reservoir engineering. In these application areas there is often a need for modelling discontinuities in the underlying data set. For geological surfaces such discontinuities are commonly known as *faults*.

In this paper the method described in [1] is generalised to situations where, in addition to the scattered data points, there are one or more given faults in the parameter domain. A fault is either an open or closed continuous planar curve. It is required that the grid function approximation is close to the values at scattered data points, is in general discontinuous across faults and is smooth elsewhere. This means in particular that values at grid nodes on one side of a fault should be largely independent of values at scattered data points lying on the other side. The method described in this paper achieves the desired discontinuity across faults and at the same time yields a smooth grid function in regions not containing faults, provided the given data is smooth.

It should be noted that the method presented here is rather specific to geological data. In a paper by Franke & Nielson [2], three ways of modelling faults were described and the problem we study falls into the second category; namely when the location of the faults is known, but the discontinuity is uknown. We have not considered situations where the location of the faults is unknown. Nor have we studied discontinuities in derivative (*creases*). Although if the data points are dense enough near a crease, our method will model the desired behaviour quite well.

The approximation scheme consists of three steps. In what follows we shall say that a point $\mathbf{q} \in \mathbb{R}^2$ is *visible* from a node $\mathbf{p} \in \Omega$ if the straight line segment from $\mathbf{p}$ to $\mathbf{q}$ does not

* This research was supported by the The Royal Norwegian Council for Scientific and Industrial Research

intersect any fault. The steps are the following:

1. Regularisation. From the given scattered data and faults determine a subset D of grid nodes in Ω which are close to at least one visible data point.
2. Local approximation. Determine a value at each node in D by interpolating nearby visible data points.
3. Extrapolation. Using the values found on nodes in D (from Step 2) and taking into account the given faults, simultaneously determine values at the remaining nodes $E = \Omega \setminus D$ of the grid using a smoothness criterion.

The extrapolation (Step 3) is greatly simplified, without apparently compromising the quality of the final approximation, by approximating the faults. Each fault is approximated by a connected set of straight line segments, each of which starts at the centre of one cell and ends at the centre of one of the four neighbouring cells. A cell is the rectangular-shaped region of $\mathbb{R}^2$ inside any 2×2 subgrid of Ω. With the faults in this form it is relatively easy to incorporate the natural boundary conditions $\partial^2/\partial n^2 = 0$ and $\partial^3/\partial n^3 = 0$ across faults since these now reduce to either $\partial^2/\partial x^2 = 0$ and $\partial^3/\partial x^3 = 0$ or $\partial^2/\partial y^2 = 0$ and $\partial^3/\partial y^3 = 0$.

In Section 2 the regularisation and local approximation are briefly described. The method is essentially the same as that described in [1]. The main difference is that the only data points taken into account at each node are those that are visible from it.

In Section 3 the extrapolation is addressed. The fault approximations and the associated smoothing operator are defined in Subsections 3.1 and 3.2.

In Subsection 3.3 it is then proved that the operator is both symmetric and, under certain mild conditions, is positive-definite. This means that the matrix equation can be uniquely solved. The conditions say essentially that there must be enough visible data and consequently that D is large enough. This is not surprising since D acts as a boundary or initial condition for the smoothing operator — every differential equation requires boundary or initial conditions in order that there should be a unique solution. The condition on D is very weak. Indeed in the case where there are no faults, it is sufficient that D contains at least four points which lie neither on a straight line nor on a hyperbola (i.e. are not simultaneously the zeros of any nonzero bilinear function). If there are faults, D must contain four points of this kind in every region of the grid which is topologically isolated from other regions by faults. This is described in greater detail towards the end of Section 3.

Finally in Section 4 numerical examples are given.

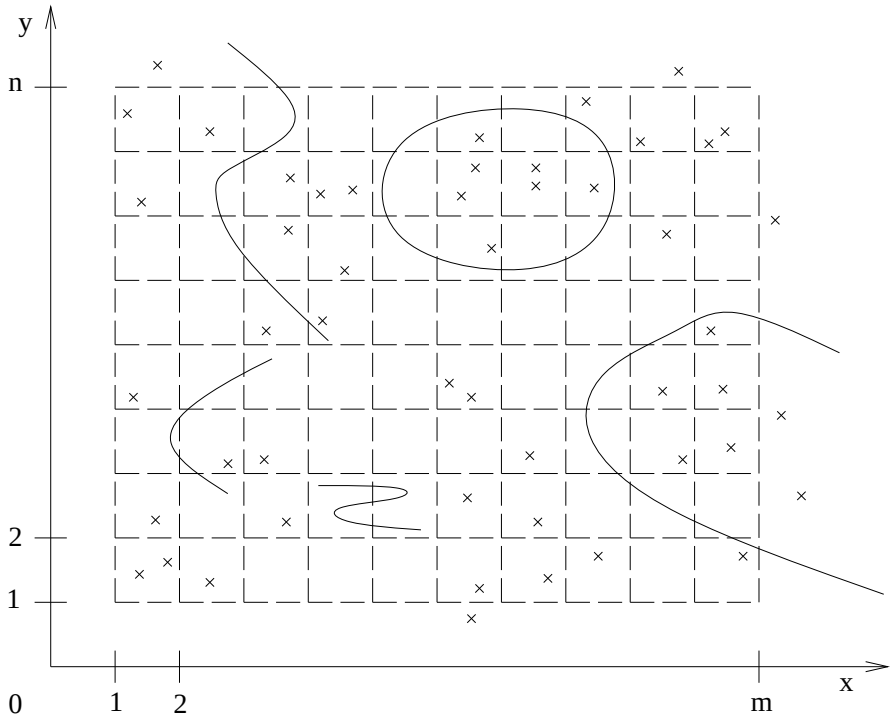


Figure 1. The faults $\mathbf{\Gamma}_k$, the scattered data points $\mathbf{p}_l$, and the grid Ω.

§2. Regularisation and local approximation

Suppose that a set of faults $\{\mathbf{\Gamma}_k\}$ is given, where each fault is a continuous planar curve $\mathbf{\Gamma}_k : [0,1] \to \mathbb{R}^2$. Suppose also that $f : \mathbb{R}^2 \to \mathbb{R}$ is some function which is smooth in any neighbourhood of $\mathbb{R}^2$ which does not intersect any fault. In general, f can have discontinuities across a fault. A sample of *scattered data* from f is a set of the form $\{(\mathbf{p}_l, z_l)\} \in \mathbb{R}^2 \times \mathbb{R}$, where $f(\mathbf{p}_l) = z_l$. Given only the faults $\{\mathbf{\Gamma}_k\}$ and the sample $\{(\mathbf{p}_l, z_l)\}$, the aim is to approximate f by a grid function g defined on some given grid lying in $\mathbb{R}^2$.

The method described in this paper can be applied to any rectangular uniform grid (and, with some reasonable constraints, any subset of such a grid) but to simplify the exposition and notation we shall assume that the grid is

$$\Omega = \{\mathbf{i} = (i, j) : 1 \le i \le m, 1 \le j \le n\}, \tag{1}$$

a subset of $\mathbb{Z}^2$. The convex hull of Ω is the rectangular region $[1, m] \times [1, n] \subset \mathbb{R}^2$. Figure 1 shows an example of scattered data points marked by crosses, faults shown as curves, and the grid Ω. The goal is to find a function $g : \Omega \to \mathbb{R}$ which has the properties:

1. the error $|f(\mathbf{i}) - g(\mathbf{i})|$ is small for each $\mathbf{i} \in \Omega$,
2. g is smooth in a discrete sense at nodes in regions not containing faults, in other words divided differences along rows and columns not crossing faults are small.

The regularisation consists of defining a subset D, referred to as a *regularly scattered data set*, of nodes in Ω where there are visible data points nearby. The local approximation consists of finding an approximation $y : D \to \mathbb{R}^2$ by using a local interpolation at each node in D. The extrapolation extends y in a smooth way to a grid function $g : \Omega \to \mathbb{R}$ defined on the whole grid.

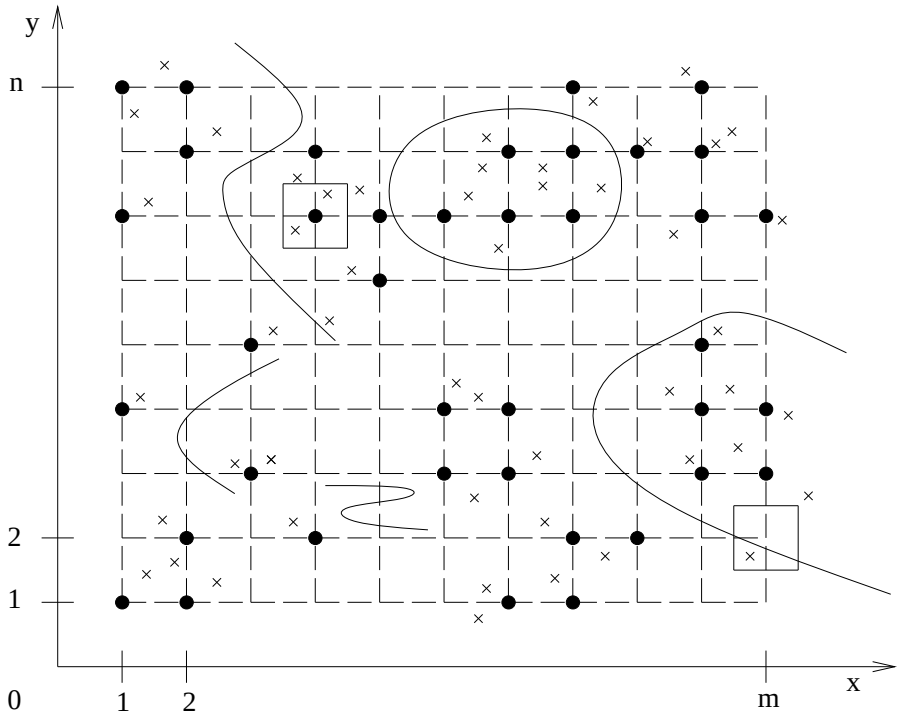


Figure 2. Regularisation — nodes belonging to D are marked with black dots.

We shall simply illustrate the first two phases of the approximation method by describing one possible method. The infinity norm is defined as

$$\|\mathbf{p}-\mathbf{q}\|_\infty = \max\{|p_1-q_1|, |p_2-q_2|\}.$$

An ϵ ball in this norm is a square of dimension $2\epsilon \times 2\epsilon$ and with sides parallel to the x and y axes. Choose two real numbers $r_1 > 0$ and $r_2 \geq r_1 > 0$.

Regularisation. Let $\mathbf{i} \in \Omega$. We say that $\mathbf{i} \in D$ if $\exists\ l$ such that $\mathbf{p}_l$ is visible from $\mathbf{i}$ (no fault intersects the line segment $\overline{\mathbf{p}_l\mathbf{i}}$) and $\|\mathbf{p}_l - \mathbf{i}\|_\infty \leq r_1$.

Local approximation. Let $\mathbf{i} \in D$. Find all points $\mathbf{p}_{l_r}$, $r = 1, \ldots, N_\mathbf{i}$, which are visible from $\mathbf{i}$ and such that $\|\mathbf{p}_{l_r} - \mathbf{i}\|_\infty \leq r_2$. There is at least one. Interpolate the data $(\mathbf{p}_{l_r}, z_{l_r})$ by a smooth function $\psi_\mathbf{i} : \mathbb{R}^2 \to \mathbb{R}$, i.e. such that $\psi_\mathbf{i}(\mathbf{p}_{l_r}) = z_{l_r}$. Then set $y_\mathbf{i} = \psi_\mathbf{i}(\mathbf{i})$.

In the numerical examples shown in this article we have chosen $\psi_\mathbf{i}$ to be a thin-plate spline, a linear combination of radial basis functions $B(r) = r^2 \log(r)$ plus linear terms; see Powell [4]. The constant r_1 was taken to be $1/2$ which means that every scattered data point is inside one and only one localising square. In Figure 2 the nodes belonging to D are marked with black dots. Two of the localising squares have been drawn. Note that the box surrounding the node $(m, 2)$ contains a point but it is not visible from $(m, 2)$. Therefore $(m, 2) \notin D$. We have also allowed r_2 to vary from node to node in such a way that at each node one always finds at least three visible points not lying in a straight line provided this is possible. This ensures that there is a unique thin-plate interpolating spline at each node $\mathbf{i} \in D$. It can, of course, happen that there are not at least three visible points. This can happen if the faults are dense around the node and the scattered data points are sparse. In this case we use a simpler local interpolation such as Shepard's method [5].

§3. Extrapolation

Some simple notation will help in defining the the extrapolation problem and its solution precisely. For any grid function $x : \mathbb{Z}^2 \to \mathbb{R}$, it is convenient to write $x_{\mathbf{j}}$ for $x(\mathbf{j})$ when $\mathbf{j} \in \mathbb{Z}^2$. For any subset $J \subset \mathbb{Z}^2$, we let $G(J)$ be the space of grid functions which are supported in J, i.e.

$$x \in G(J) \Leftrightarrow x_{\mathbf{j}} = 0 \quad \forall \mathbf{j} \notin J.$$

Using this notation the extrapolation problem can be phrased in the following way.

Extrapolation problem. Given $y \in G(D)$, find $x \in G(E)$ such that $g = x + y$ is an extension of y which is smooth in regions not containing faults.

In order to define the solution to the extrapolation problem we need some notation connected to linear grid point operators.

A linear operator $Q : G(\mathbb{Z}^2) \to G(\mathbb{Z}^2)$ can be represented by a *mask* $Q = (Q_{\mathbf{i}}(\mathbf{j}))_{\mathbf{i},\mathbf{j} \in \mathbb{Z}^2}$ which operates on an element $x \in G(\mathbb{Z}^2)$ as a discrete convolution

$$(Qx)_{\mathbf{j}} = \sum_{\mathbf{i} \in \mathbb{Z}^2} Q_{\mathbf{i}}(\mathbf{j}) x_{\mathbf{j}-\mathbf{i}}. \tag{2}$$

We note that we allow the mask Q to vary over the domain of interest, thus the mask connected to a given $\mathbf{j} \in \mathbb{Z}^2$ is denoted $Q(\mathbf{j})$.

We will solve the extrapolation problem by using a linear smoothing operator $P_F : G(\Omega) \to G(E)$ which is a discretisation of the fourth order differential operator $\partial^4/\partial x^4 + \partial^4/\partial y^4$, with the natural boundary conditions $\partial^2/\partial n^2 = \partial^3/\partial n^3 = 0$ across faults. P_F gives a value at every node $\mathbf{i} \in E$. Our solution to the extrapolation problem is then

Solution to the extrapolation problem. Let $x \in G(E)$ be the solution of the linear system

$$(P_F x)_{\mathbf{i}} = -(P_F y)_{\mathbf{i}} \qquad \text{for } \mathbf{i} \in E. \tag{3}$$

Then $g = x + y$ is the required grid function approximation.

The resulting grid approximation g must then have the property that at every node $\mathbf{i} \in E$, $(P_F g)_{\mathbf{i}} = 0$. This implies that g minimises a discretisation of $g_{xx}^2 + g_{yy}^2$, a measure of roughness, implying that g smooth in E. It will be shown that under certain (very weak) conditions on the set D, that x exists and is unique. Essentially there merely need to be enough nodes in D since the values at nodes in D act as boundary conditions for the operator equation. The solvability is proved by showing that P_F as an operator $G(E) \to G(E)$ is both symmetric and positive definite.

The operator P_F is based on the operator P defined below which is the usual discretisation of $\partial^4/\partial x^4 + \partial^4/\partial y^4$. Let the basis vectors be $\mathbf{d}^1 = (1, 0)$ and $\mathbf{d}^2 = (0, 1)$. Then $P : G(\mathbb{Z}^2) \to G(\mathbb{Z}^2)$ is the operator

$$\begin{aligned} (Px)_{\mathbf{j}} = \quad & 12x_{\mathbf{j}} \\ & - 4[x_{\mathbf{j}-\mathbf{d}^1} + x_{\mathbf{j}+\mathbf{d}^1} + x_{\mathbf{j}-\mathbf{d}^2} + x_{\mathbf{j}+\mathbf{d}^2}] \\ & + [x_{\mathbf{j}-2\mathbf{d}^1} + x_{\mathbf{j}+2\mathbf{d}^1} + x_{\mathbf{j}-2\mathbf{d}^2} + x_{\mathbf{j}+2\mathbf{d}^2}], \end{aligned} \tag{4}$$

for $\mathbf{j} \in \mathbb{Z}^2$. It will be helpful to split this operator up into the two directions so let $P = H + V$ where

$$\begin{aligned} (Hx)_{\mathbf{j}} =& x_{\mathbf{j}-2\mathbf{d}^1} - 4x_{\mathbf{j}-\mathbf{d}^1} + 6x_{\mathbf{j}} - 4x_{\mathbf{j}+\mathbf{d}^1} + x_{\mathbf{j}+2\mathbf{d}^1}, \\ (Vx)_{\mathbf{j}} =& x_{\mathbf{j}-2\mathbf{d}^2} - 4x_{\mathbf{j}-\mathbf{d}^2} + 6x_{\mathbf{j}} - 4x_{\mathbf{j}+\mathbf{d}^2} + x_{\mathbf{j}+2\mathbf{d}^2}. \end{aligned} \tag{5}$$

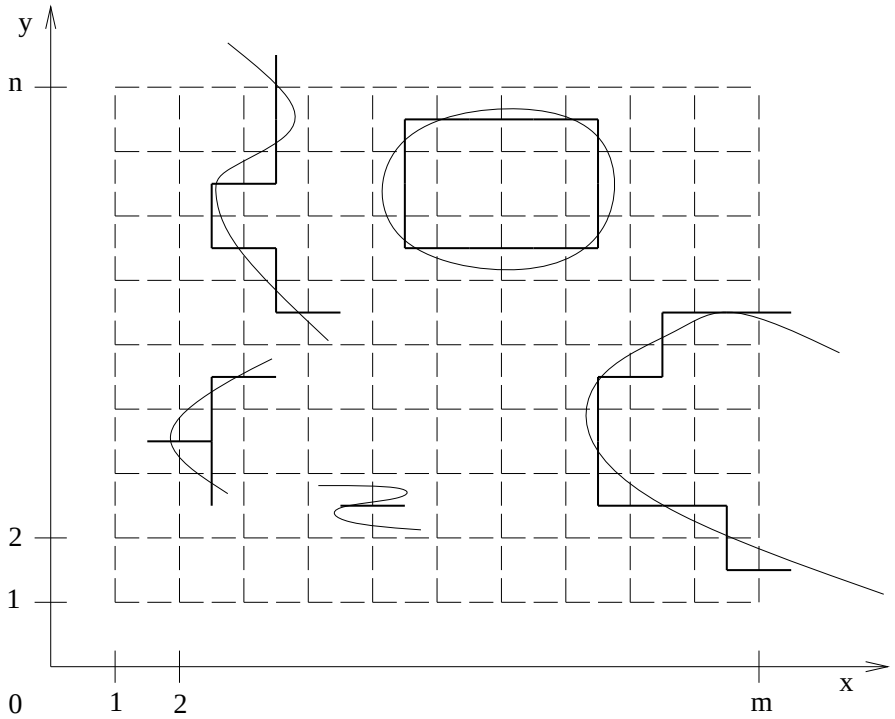


Figure 3. Approximation of the faults by line segments.

We shall define P_F to be an adaptation of P adjusted according to faults in the grid. In particular, the mask is adjusted at any node close to a fault in order to obtain the natural boundary conditions across each fault, i.e. that the second and third normal derivatives are zero.

3.1 Faults.

The basic idea behind our treatment of faults is to approximate them by sets of line segments, called *fault segments*. A fault segment is a line segment parallel with either the x or y axis starting at the centre of one cell and ending at the centre of an adjacent cell. Though each fault segment is a line segment it is more convenient to represent it by its midpoint. Then a fault segment is a point either of the form $\mathbf{i}+\mathbf{d}^1/2$ or $\mathbf{i}+\mathbf{d}^2/2$. Each fault (curve) $\mathbf{\Gamma}_k$ is approximated by set of vertical and horizontal fault segments.

Definition 3.1. The approximation of $\mathbf{\Gamma}_k$ is the the union $F_{k,v} \cup F_{k,h}$ where

$$F_{k,v} = \{\mathbf{i}+\mathbf{d}^1/2 : \mathbf{i}=(i,j), 1 \le i \le m-1, 1 \le j \le n, \mathbf{\Gamma}_k \cap \overline{\mathbf{i}\ \mathbf{i}+\mathbf{d}^1} \neq \emptyset\},$$
$$F_{k,h} = \{\mathbf{i}+\mathbf{d}^2/2 : \mathbf{i}=(i,j), 1 \le i \le m, 1 \le j \le n-1, \mathbf{\Gamma}_k \cap \overline{\mathbf{i}\ \mathbf{i}+\mathbf{d}^2} \neq \emptyset\}.$$

It is clear that because each fault is continuous, the set of fault segments approximating it is connected. In general though, the approximation is not a single continuous curve, it may have a tree-like structure. More than two fault segments can meet at the centre of a cell; see Figure 3.

The approximation of a fault means that when a fault crosses a vertical grid line, the derivative of f normal to the fault at that point is approximated by the partial derivative f_y at the point midway between the two adjacent nodes. Thus there are really two things that are approximated — the intersection point and the direction normal to the fault at that point. Despite this, the numerical examples we have run give excellent results. This

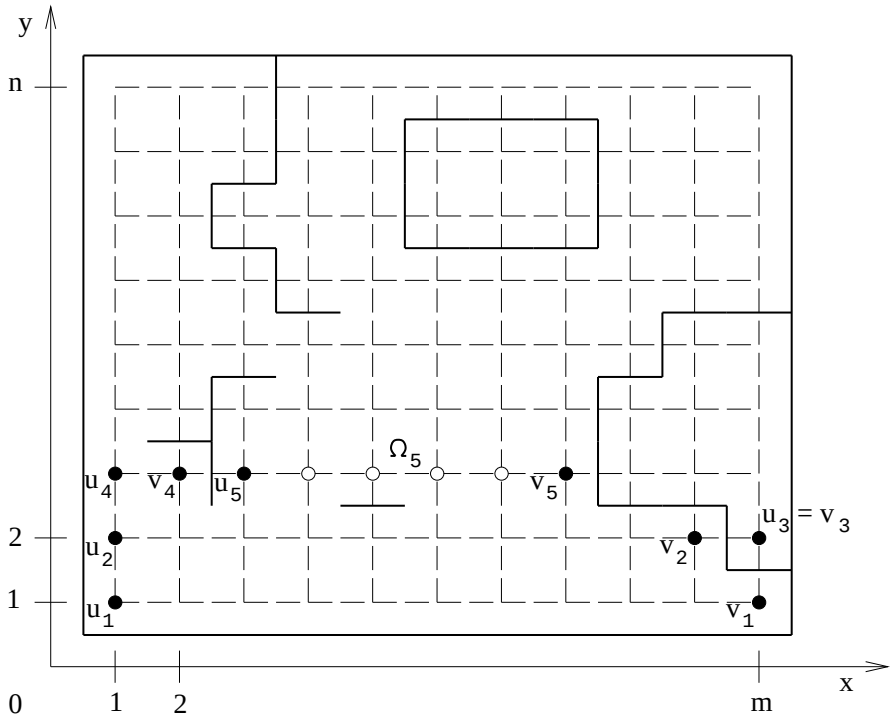


Figure 4. The set of faults segments F together with some of the nodes $\mathbf{u}_r$, $\mathbf{v}_r$ and subdomains Ω_r.

is perhaps due to the fact that the approximation is only defined at the nodes of a discrete grid. It is not so important *where* a fault crosses a grid line between two nodes, only *whether* it does so.

In order to introduce natural boundary conditions at the edge of the grid we add fault segments to the boundary. Thus the whole set of fault segments consists of the union of all the $F_{k,v}$'s and $F_{k,h}$'s and the boundary segments.

Definition 3.2. Let Ω be given by (1) and let $\mathbf{\Gamma}_k$ be defined as earlier. Then the set F of fault segments associated with Ω and $\{\mathbf{\Gamma}_k\}$ is the set

$$F = F_v \cup F_h, \tag{6}$$

where

$$F_v = \bigcup_k F_{k,v} \quad \cup \quad \{(1/2, j) : 1 \le j \le n\} \quad \cup \quad \{(m + 1/2, j) : 1 \le j \le n\},$$

$$F_h = \bigcup_k F_{k,h} \quad \cup \quad \{(i, 1/2) : 1 \le i \le m\} \quad \cup \quad \{(i, n + 1/2) : 1 \le i \le m\}.$$

The set F_h is called the set of *horizontal fault segments* while F_v is called the set of *vertical fault segments*. Figure 4 shows the complete set of fault segments, including the outer boundary in our example.

3.2 Definition of the fault operator.

The definition of the fault operator P_F relies on the way the operators V and H are adjusted in regions close to horizontal and vertical faults respectively. In order to define

H_F, consider F_v defined in (6). According to the vertical fault segments we can uniquely partition Ω into N disjoint subdomains $\Omega_1, \ldots, \Omega_N$ of the form

$$\Omega_r = \{\mathbf{i} : \mathbf{i} = \mathbf{u}_r, \mathbf{u}_r + \mathbf{d}^1, \mathbf{u}_r + 2\mathbf{d}^1, \ldots, \mathbf{v}_r\}$$

where $\mathbf{u}_r - \mathbf{d}^1/2$ and $\mathbf{v}_r + \mathbf{d}^1/2$ for $r = 1, \ldots, N$ are consecutive vertical faults. Figure 4 illustrates the fault segments and some of the subdomains Ω_r associated with the vertical faults. Based on the definition of these subdomains, we can now define the horizontal part H_F of the fault operator P_F.

Definition 3.3. Let $H : G(\mathbb{Z}^2) \to G(\mathbb{Z}^2)$ be defined in (5), and let the set of faults F be defined as in (6). Then we define $H_F : G(\mathbb{Z}^2) \to G(\mathbb{Z}^2)$ by

$$(H_F x)_{\mathbf{i}} = \begin{cases} (H h_r(x))_{\mathbf{i}} & \text{when } \mathbf{i} \in \Omega_r \text{ for some } r; \\ 0 & \text{otherwise} \end{cases}$$

where the mapping $h_r : G(\mathbb{Z}^2) \to G(\mathbb{Z}^2)$ is defined by

$$h_r(x)_{\mathbf{i}} = \begin{cases} x_{\mathbf{i}} & \text{when } \mathbf{i} \in \Omega_r, \\ 2x_{\mathbf{u}_r} - x_{\mathbf{u}_r + \mathbf{d}^1} & \text{when } \mathbf{i} = \mathbf{u}_r - \mathbf{d}^1, \\ 3x_{\mathbf{u}_r} - 2x_{\mathbf{u}_r + \mathbf{d}^1} & \text{when } \mathbf{i} = \mathbf{u}_r - 2\mathbf{d}^1, \\ 2x_{\mathbf{v}_r} - x_{\mathbf{v}_r - \mathbf{d}^1} & \text{when } \mathbf{i} = \mathbf{v}_r + \mathbf{d}^1, \\ 3x_{\mathbf{v}_r} - 2x_{\mathbf{v}_r - \mathbf{d}^1} & \text{when } \mathbf{i} = \mathbf{v}_r + 2\mathbf{d}^1, \\ 0 & \text{otherwise,} \end{cases}$$

if $|\Omega_r| > 1$. If $|\Omega_r| = 1$ define $h_r = 0$. ■

In a similar way one can partition Ω according to the horizontal faults in each column. One can therefore define the operator V_F in a completely analogous way to H_F. There is no point defining this formally since all arguments below will be symmetrical with respect to H_F and V_F. Finally we set $P_F = H_F + V_F$.

The definition of h_r represents a discretisation of the natural boundary conditions $u_{xx} = u_{xxx} = 0$ at $\mathbf{u}_r - \mathbf{d}^1/2$ and $\mathbf{v}_r + \mathbf{d}^1/2$. Indeed, let $H^{\frac{1}{2}} : G(\mathbb{Z}^2) \to G(\mathbb{Z}^2)$ denote the second order difference operator defined by

$$(H^{\frac{1}{2}} x)_{\mathbf{i}} = x_{\mathbf{i} - \mathbf{d}^1} - 2x_{\mathbf{i}} + x_{\mathbf{i} + \mathbf{d}^1}. \tag{7}$$

Then we find from the definition of h_r that

$$(H^{\frac{1}{2}} h_r(x))_{\mathbf{i}} = 0 \tag{8}$$

for $\mathbf{i} \in \{\mathbf{u}_r - \mathbf{d}^1, \mathbf{u}_r, \mathbf{v}_r, \mathbf{v}_r + \mathbf{d}^1\}$.

We will consider symmetry and positive definite properties with respect to the usual Euclidian inner product on $G(\mathbb{Z}^2)$ defined by

$$< x, y >= \sum_{\mathbf{i} \in \mathbb{Z}^2} x_{\mathbf{i}} y_{\mathbf{i}}.$$

The following lemma is the first step on this path.

Lemma 3.4. *Let $H^{\frac{1}{2}}$ be given by (7). If $x \in G(\Omega_s)$ and $y \in G(\Omega_r)$ we have*

$$< H_F x, y >= \begin{cases} \sum_{\mathbf{i}\in\Omega_r\setminus\{\mathbf{u}_r,\mathbf{v}_r\}} (H^{\frac{1}{2}}x)_{\mathbf{i}}(H^{\frac{1}{2}}y))_{\mathbf{i}} & \text{if } r = s; \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Proof. First note that by the definition of H_F and since y vanishes outside Ω_r we have

$$\begin{aligned} < H_F x, y > &= \sum_{\mathbf{i}\in\Omega_r} (H_F x)_{\mathbf{i}} y_{\mathbf{i}} \\ &= \sum_{\mathbf{i}\in\Omega_r} (H h_r(x))_{\mathbf{i}} y_{\mathbf{i}} \\ &=< H h_r(x), y > \end{aligned} \tag{10}$$

Now, by the definition of h_r it follows that $h_r(x) = 0$ if $s \neq r$. Therefore, the second equality in (10) implies that $< H_F x, y >= 0$ unless $r = s$. This proves (9) for $s \neq r$.

Assume that $s = r$. Then the remaining part follows from the following string of equalities:

$$\begin{aligned} < H h_r(x), y > &=< H^{\frac{1}{2}} h_r(x), H^{\frac{1}{2}} y > \\ &= \sum_{\mathbf{i}\in\mathbb{Z}^2} (H^{\frac{1}{2}} h_r(x))_{\mathbf{i}} (H^{\frac{1}{2}} y)_{\mathbf{i}} \\ &= \sum_{\mathbf{i}\in\Omega_r\setminus\{\mathbf{u}_r,\mathbf{v}_r\}} (H^{\frac{1}{2}} h_r(x))_{\mathbf{i}} (H^{\frac{1}{2}} y)_{\mathbf{i}} \\ &= \sum_{\mathbf{i}\in\Omega_r\setminus\{\mathbf{u}_r,\mathbf{v}_r\}} (H^{\frac{1}{2}} x)_{\mathbf{i}} (H^{\frac{1}{2}} y)_{\mathbf{i}}. \end{aligned}$$

Here, the first equality follows by using the fact that $< H\phi, \psi >=< H^{\frac{1}{2}}\phi, H^{\frac{1}{2}}\psi >$ for any $\phi, \psi \in G(\mathbb{Z}^2)$ of finite support. The third equality follows by (8) and by using the fact that y vanishes outside Ω_r. The fourth follows since h_r restricted to $G(\Omega_r)$ is the identity mapping. ■

3.3 Symmetry and positive definiteness.

In this section we will show that the fault operator P_F is symmetric on $G(\Omega)$, and derive sufficient conditions on the subdomains D and E relative to the fault set F to ensure that P_F is positive definite on $G(E)$.

Symmetry and positive semi-definiteness are properties not related to the given data set and are a consequence of (9) as shown in the following theorem.

Theorem 3.5. *P_F is symmetric and positive semi-definite on $G(\Omega)$.*

Proof. It is required to demonstrate that $< P_F x, y >=< P_F y, x >$ and $< P_F x, x >\geq 0$ for any $x, y \in G(\Omega)$. Now, since $< P_F x, y >=< H_F x, y > + < V_F x, y >$ and since the arguments for H_F and V_F are analogous we only argue for H_F.

Let x^r and y^r be the projections of x and y onto $G(\Omega_r)$ for $r = 1, \ldots, N$. Then by linearity

$$< H_F x, y >= \sum_{p=1}^{N}\sum_{q=1}^{N} < H_F x^p, y^q > .$$

The assertion then follows by a direct application of (9). ■

In order to guarantee strict positive definiteness on the subspace $G(E)$, we have to impose conditions on the subdomains D and E relative to the fault segments F. These conditions guarantee that the linear system (3) is non-singular. Moreover, they make the conjugate-gradient method applicable [3].

In order to define these conditions we need a concept of connectedness. The main point is that the fault segments may split the nodes in Ω up into a number of isolated subsets. This will certainly be the case if the faults $\mathbf{\Gamma}_k$ split the rectangle $[1, m] \times [1, n]$ up into disconnected regions. For this reason, we decompose Ω into connected components bounded by fault segments and then the condition for positive-definiteness is defined in each component.

The ideas for the definition of connectedness and the lemma concerning the unique decomposition of a set into connected components are taken from topology; see Simmons [6]. What is special here is that the meaning of connectedness depends on the size of the mask. From the point of view of the operator P_F, two nodes are not in the same connected set unless they belong to a path of complete cells, where a cell is *complete* if it does not intersect any fault segments. Indeed, due to the size of the mask of P_F, a subset of Ω can be thought of as having 'measure zero' unless it contains at least one complete cell. If we had used a lower order operator such as the Laplacian operator, $\partial^2/\partial x^2 + \partial^2/\partial y^2$, any non-empty set of nodes would have had non-zero measure.

Definition 3.4.

(i) A *cell* c is a subset of $\mathbb{R}^2$ of the form $[i, i+1] \times [j, j+1]$ where $1 \leq i < m$ and $1 \leq j < n$ are integers.

(ii) A cell c is *complete* if $c \cap F = \emptyset$ where F is the given set of faults defined in (6).

(iii) The set of all complete cells is denoted by C.

(iv) A *path* from $c \in C$ to $d \in C$ is a finite sequence of complete cells $c_0, \ldots, c_M \in C$ such that $c_0 = c$, $c_M = d$ and such that $c_{\mu-1}$ and c_μ share a common edge (not merely a node) for $\mu = 1, \ldots, M$.

(v) A subset $A \subset C$ is *connected* if $\forall c, d \in A$, there exists a path from c to d which is contained in A.

(vi) A connected subset $A \subset C$ is a *component* of C if for every connected set $B \subset C$ with $A \cap B \neq \emptyset$, we have $B \subset A$, i.e. A is a maximal connected set.

Having made the right definitions it is now relatively straightforward to show that the set C can be uniquely decomposed into components, a well-known result for topological spaces. The situation here is a little different to that in topology because we have defined connectness without defining a topology.

Lemma 3.6. *Every cell $c \in C$ belongs to exactly one component of C.*

Proof. Let $c \in C$ and let $\{A_i\}$ be the class of all connected subsets of C which contain c. This class is non-empty since $\{c\}$ is connected. Since c is a member of all the A_i's it is easy to see that also $A = \cup_i A_i$ is connected. Indeed one can construct a path contained in A between any two cells d_1 and d_2 in A consisting of two subpaths; one from d_1 to c and the other from c to d_2. Now let B be any connected set with $A \cap B \neq \emptyset$, and let $b \in A \cap B$. Let $\{c_\mu\} \subset A$ be a path from b to c. Then $B \cup \{c_\mu\}$ is a connected set containing c and hence $B \cup \{c_\mu\} \subset A$. This means that also $B \subset A$, and hence A is a component of C. This proves that c is a member of the component A.

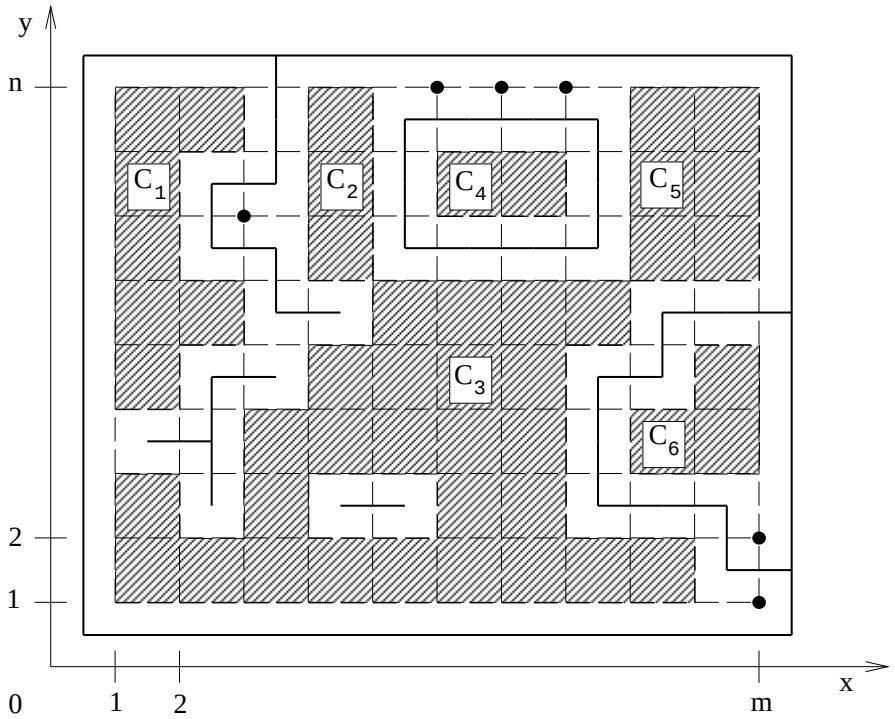


Figure 5. The connected components of C. Nodes not contained in $[C]$ are marked by black dots.

For uniqueness of the component suppose that c is a member of a component A'. Then, because A' is connected, it is clearly among the A_i's and therefore $A' \subset A$. But since A' is also a component of C and $A' \cap A \neq \emptyset$, we must also have $A \subset A'$. Therefore $A' = A$. ■

A corollary of the lemma is that we can decompose C uniquely into its disjoint components, in other words

$$C = \bigcup_{\alpha=1}^{R} C_\alpha, \qquad \text{and} \qquad \alpha \neq \beta \Rightarrow C_\alpha \cap C_\beta = \emptyset.$$

There are six components in our example shown in Figure 5.

If A is any set of cells let $[A]$ denote the set of all *nodes* contained in the cells in A. Note that it is not necessarily the case that $[C_\alpha] \cap [C_\beta] = \emptyset$ since two components can touch at a single corner; this is the case for $[C_2] \cap [C_3]$ in Figure 5. Usually $[C]$ will be either equal to Ω or be a very large subset of it; we may regard the set $\Omega \backslash [C]$ as having measure zero.

We are now in position to address the positive definiteness of the fault operator P_F on the sub-space $G(E) \subset G(\Omega)$. The following lemma solves the main part of the problem.

Lemma 3.7. *Let $x \in G(\Omega)$ and suppose that $< P_F x, x >= 0$. Then the restriction of x to any set $[C_\alpha]$ is bilinear.*

Proof. Note first that by (7) and (9) we have

$$< H_F x, x >= \sum_{r=1}^{N} \sum_{\mathbf{i} \in \Omega_r \backslash \{\mathbf{u}_r, \mathbf{v}_r\}} (x_{\mathbf{i}-\mathbf{d}^1} - 2x_{\mathbf{i}} + x_{\mathbf{i}+\mathbf{d}^1})^2.$$

By rewriting the sum one finds

$$< H_F x, x >= \sum_{\substack{\mathbf{i}\in\Omega \\ \mathbf{i}\pm\mathbf{d}^1/2\notin F_v}} (x_{\mathbf{i}-\mathbf{d}^1} - 2x_{\mathbf{i}} + x_{\mathbf{i}+\mathbf{d}^1})^2.$$

Now, since $< P_F x, x >=< H_F x, x > + < V_F x, x >$ a similar argument for V_F gives us

$$< P_F x, x >= \sum_{\substack{\mathbf{i}\in\Omega \\ \mathbf{i}\pm\mathbf{d}^1/2\notin F_v}} (x_{\mathbf{i}-\mathbf{d}^1} - 2x_{\mathbf{i}} + x_{\mathbf{i}+\mathbf{d}^1})^2 + \sum_{\substack{\mathbf{i}\in\Omega \\ \mathbf{i}\pm\mathbf{d}^2/2\notin F_h}} (x_{\mathbf{i}-\mathbf{d}^2} - 2x_{\mathbf{i}} + x_{\mathbf{i}+\mathbf{d}^2})^2. \qquad (11)$$

Assume that $< P_F x, x >= 0$. Then (11) implies that

$$\begin{aligned} x_{\mathbf{i}-\mathbf{d}^1} - 2x_{\mathbf{i}} - x_{\mathbf{i}+\mathbf{d}^1} = 0 \quad \forall \mathbf{i} \in \Omega : \mathbf{i} \pm \mathbf{d}^1/2 \notin F_v \\ x_{\mathbf{i}-\mathbf{d}^2} - xy_{\mathbf{i}} - x_{\mathbf{i}+\mathbf{d}^2} = 0 \quad \forall \mathbf{i} \in \Omega : \mathbf{i} \pm \mathbf{d}^2/2 \notin F_h \end{aligned} \qquad (12)$$

Choose any component C_α and any cell $c \in C_\alpha$. Then there is a unique bilinear grid function $y \in G(\mathbb{Z}^2)$, agreeing with x at the nodes at the corners of c. Moreover, since y is bilinear we have

$$\begin{aligned} y_{\mathbf{i}-\mathbf{d}^1} - 2y_{\mathbf{i}} + y_{\mathbf{i}+\mathbf{d}^1} = 0 \\ y_{\mathbf{i}-\mathbf{d}^2} - 2y_{\mathbf{i}} + y_{\mathbf{i}+\mathbf{d}^2} = 0 \end{aligned} \qquad (13)$$

for any $\mathbf{i} \in \mathbb{Z}^2$.

Now, if $d \in C_\alpha$ shares a common edge with c, then (12) and (13) implies that $y_{\mathbf{i}} = x_{\mathbf{i}}$ for all $\mathbf{i} \in c \cup d$. By induction, this implies that $y_{\mathbf{i}} = x_{\mathbf{i}}$ for all $\mathbf{i} \in \cup_\mu d_\mu$ for any path $\{d_\mu\}$ extending from c. Since any cell in C_α can be reached by a path from c and any $\mathbf{j} \in [C_\alpha]$ is in some cell in C_α, we must have $x|[C_\alpha] = y|[C_\alpha]$. Therefore x is bilinear in $[C_\alpha]$ as claimed. ■

Using Lemma 3.7, a condition on D for P_F to be positive-definite is given in the following theorem.

Theorem 3.8. *Let Ω be defined in (1) and let F be a set of faults as defined in (6). Further let C be the set of all complete cells in Ω, and let $C_1, \ldots, C_R$ be the decomposition of C into its disjoint components. If*

(i) each set $[C_\alpha] \cap D$ contains four nodes which are not the zeros of any nonzero bilinear function $f(u, v) = a + bu + cv + duv$,

and

(ii) $[C] = \Omega$

then P_F is positive definite on $G(E)$.

Proof. Suppose that $x \in G(E)$ and $< P_F x, x >= 0$. It is required to show that $x = 0$. By (ii) it is sufficient to show that $x|[C] = 0$. This is done by showing that $x|[C_\alpha] = 0$ for all α.

To this end choose a component C_α. Lemma 3.7 clearly applies and we find that the restriction of x to $[C_\alpha]$ is bilinear. Now, by condition (i), there exist four nodes $\mathbf{i}_1$, $\mathbf{i}_2$, $\mathbf{i}_3$, $\mathbf{i}_4$ in $D \cap [C_\alpha]$ which are not the zeros of any nonzero bilinear function. Therefore there is a unique bilinear function $y \in G(\mathbb{Z}^2)$ which agrees with x at these four nodes, i.e. such that

$y_{\mathbf{i}_\mu} = x_{\mathbf{i}_\mu}$, $\mu = 1, 2, 3, 4$. But since $x \in G(E)$, $x_{\mathbf{i}_\mu} = 0$ and therefore y must be the zero function. As a consequence, the uniqueness of y implies that $x|[C_\alpha] = 0$. ■

In the example in Figure 5 one can see, by comparing with Figure 2, that there are at least four nodes in each $[C_\alpha]$ which belong to D. Moreover there are at least four which do not lie on either a straight line or a hyperbola with asymptotes parallel to the x and y axes — the condition for a bilinear interpolation to be non-singular. Therefore condition (i) holds. On the other hand, one can see that $\Omega\backslash[C]$ is non-empty — these nodes are marked with black dots. Therefore condition (ii) of the theorem does not hold.

In practice the simplest solution to this problem is to isolate any nodes in $\Omega\backslash[C]$ by extra fault segments. One can augment the fault set to

$$F' = F \cup \{\mathbf{i} \pm \mathbf{d}^1/2 : \mathbf{i} \in [C], \mathbf{i} \pm \mathbf{d}^1 \in \Omega\backslash[C]\} \cup \{\mathbf{i} \pm \mathbf{d}^2/2 : \mathbf{i} \in [C], \mathbf{i} \pm \mathbf{d}^2 \in \Omega\backslash[C]\}, \quad (14)$$

and define $P_{F'}$ accordingly. The new fault segments put a boundary between the subsets $[C]$ and $\Omega\backslash[C]$. It is clear that the components C_α are unaffected by this. Then, instead of solving (3) on the whole of E, we restrict the problem to $E \cap [C]$ and it is easy to deduce the following theorem.

Theorem 3.9. *Again let Ω be defined in (1) and let F' be the augmented fault set (14). If each set $[C_\alpha] \cap D$ contains four nodes which are not the zeros of any nonzero bilinear function, then $P_{F'}$ is positive definite on $G(E \cap [C])$.*

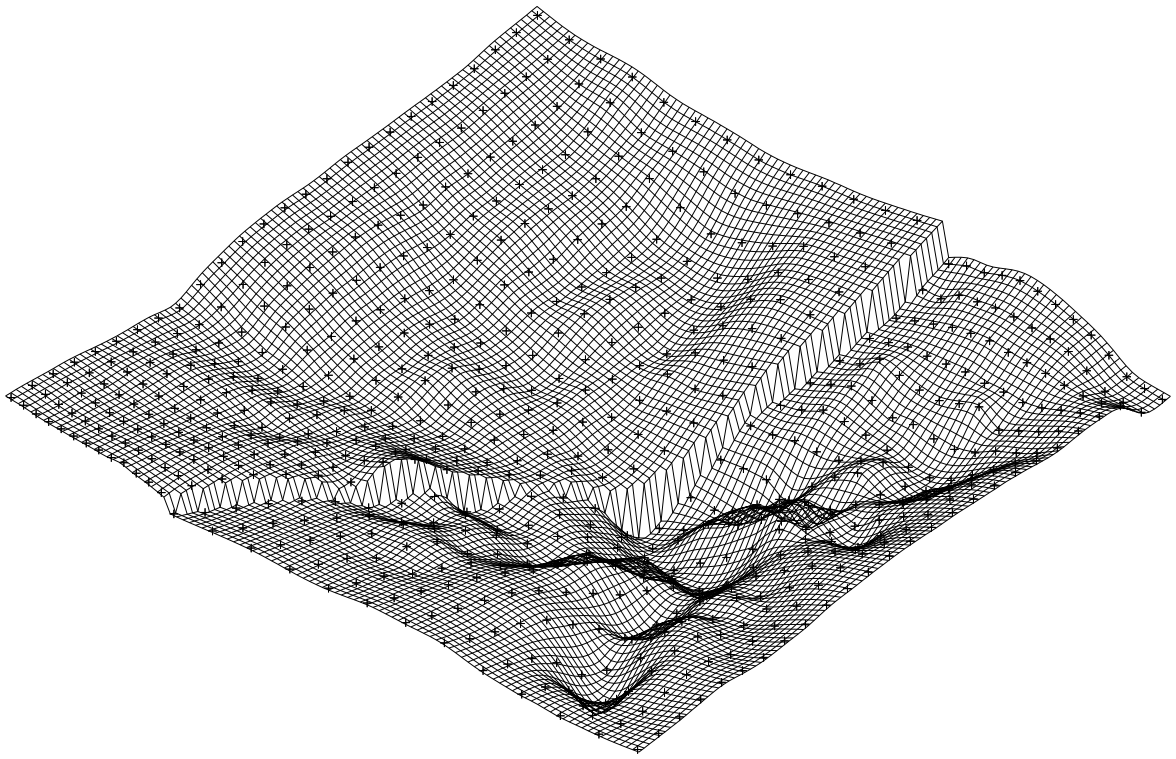
Proof. If $x \in G(E \cap [C])$ and $< P_{F'}(x), x >= 0$, then from the condition on $[C_\alpha] \cap D$, one shows that $x|[C] = 0$ using the same arguments as in Theorem 3.8 with F replaced by F'. But then $x|(E \cap [C]) = 0$. ■

Values at any nodes in $E \cap (\Omega\backslash[C])$ can be computed *afterwards* by some other means. In the numerical examples we ran, we achieved this by reversing the extrapolation. We used the *Laplacian* mask (disregarding faults) over the whole grid, regarding $G(E')$ as the unknown and $G(D')$ as data where

$$E' = E \cap (\Omega\backslash[C]), \qquad \text{and} \qquad D' = \Omega\backslash E'.$$

It is even possible to improve on this method so as to incorporate as many nodes as possible in $E \cap (\Omega\backslash[C])$ in the main extrapolation. A comparison of Figure 5 with Figure 2 shows that $E \cap (\Omega\backslash[C])$ is non-empty in our example although this example is rather extreme since the grid is so coarse. Instead of removing all nodes in $\Omega\backslash[C]$ from the extrapolation, one could remove only those which cannot be connected to one of the components $[C_\alpha]$ along a grid line not intersecting F, or even a sequence of such lines. In Figure 5, all nodes in $[C_\alpha]$ can be included in this way and, in fact, the original operator P_F is positive definite on $G(E)$ in this example. However in theory, there are much more extreme cases and the advantage of Theorem 3.9 is that it avoids the need to tackle abnormalities of this kind. In conclusion, it is possible to weaken condition (ii) in Theorem 3.8 but at the cost of extending the algorithm considerably.

Checking positive-definiteness. One can determine the connected components of Ω by running through all the cells, flagging them once they are visited. If a cell is complete, one finds the component containing it by recursively growing outwards in horizontal or vertical



Example 1. Grid approximation of seismic data and a single open fault.

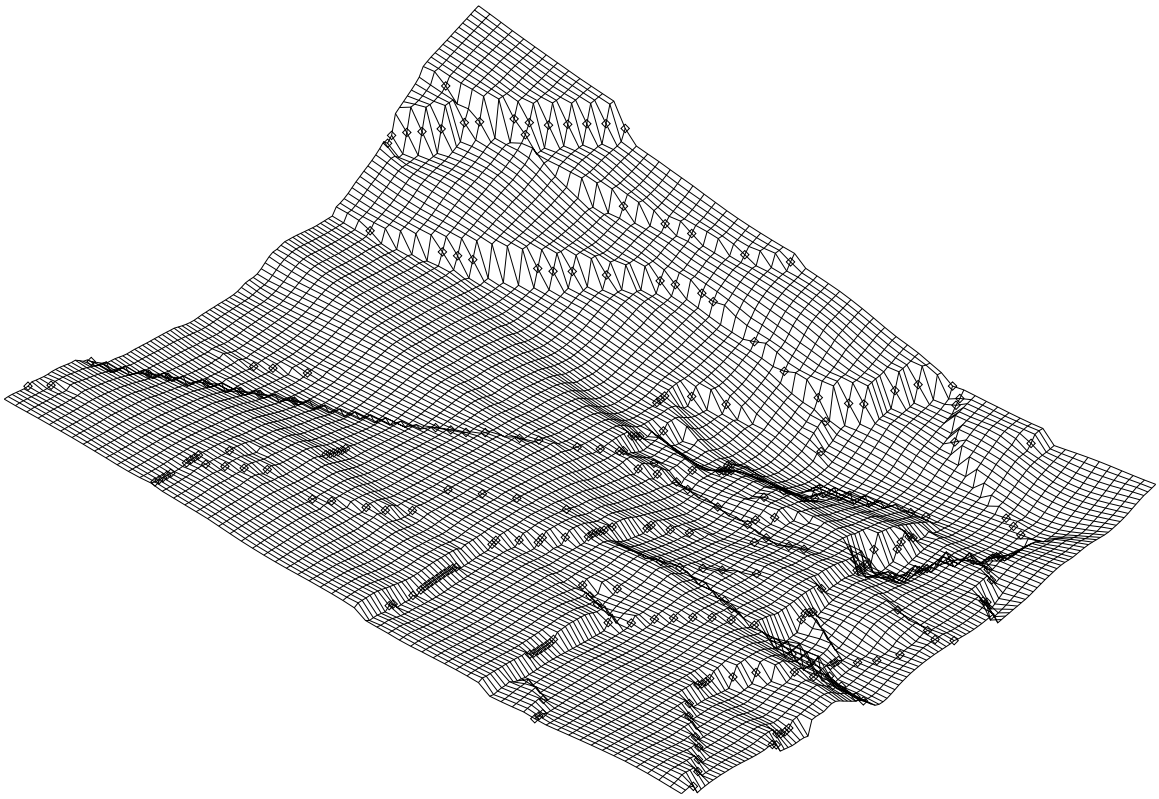
steps. Once a component is determined, the member cells are all flagged as visited, and the search continues for other components. In this way the decomposition of Ω is obtained.

Now in each component $[C_\alpha]$, one can loop through every combination of four nodes in $[C_\alpha] \cap D$ until four are found which are not the zeros of any non-zero bilinear function. This condition is equivalent to the non-singularity of a 4×4 matrix which can be determined very easily from a singular value decomposition. Often in practice, the first four nodes fulfil the criterion.

§4. Numerical Examples

In the numerical examples presented below, the conjugate-gradient method was used to solve the linear system (3). The advantage of this iterative method is that it is not, in fact, necessary to solve the system exactly. One can stop the iteration when the error between the next step and the previous one is negligible.

Example 1. The data in the first example is taken from seismic data collected from an area of the North Sea. The scattered data $\{(\mathbf{p}_l, z_l)\}_l$ are shown as crosses and there is just one fault $\mathbf{\Gamma}_1$ running across the domain (not shown). The scattered data here is quite regularly



Example 2. Grid approximation of seismic scattered data and closed faults.

scattered. The computed grid function $g : \Omega \to \mathbb{R}^2$ is shown where values at adjacent nodes are joined by line segments. The grid Ω is split into precisely two components by the fault and the conditions for Theorem 3.8 hold ($\Omega = [C] = [C_1] \cup [C_2]$). Note that the grid function is smooth everywhere except across the fault.

Example 2. The data in the second example is a section of seismic data taken from a layer of an oil field in the North Sea. The grid shown is only a subset of the whole 181×361 grid which was computed by our method. The scattered data, in the form of contours, is not shown as it would make the figure very unclear.

There are many faults in this example and they are all closed curves, generally long and thin. One can see where the faults are from the jumps in the grid function and it is noticeable that Ω is split into many isolated components. Those nodes of the grid which lie inside one of the faults are *deliberately* removed from the main extrapolation in the same way as those in $\Omega \backslash [C]$ by augmenting the fault set. All such nodes are marked with boxes. The values of the grid function at these nodes were computed by applying the discrete Laplacian operator after the other values had been calculated. The effect of this is that g is bilinear in nature inside closed faults and interpolates between two sides of a fault region. This is the desirable

effect in seismic applications.

§5. References

1. E. Arge, M. Dæhlen, A. Tveito, A strategy for the approximation of large scattered data sets, to appear in J. Comp. Appl. Math., 1992.
2. R. Franke, G. Nielson, Surface approximation with imposed conditions, in Surfaces in Computer Aided Design, R. Barnhill and W. Boehm (eds) North Holland, Amsterdam, 1983, 135–146.
3. G. H. Golub, C. F. Van Loan, Matrix Computations, Johns Hopkins University Press, 1989.
4. M. J. D. Powell, The theory of radial basis function approximation in 1990, Advances in Numerical Analysis, Vol. II, W. Light (ed.), Oxford Science Publications, 1992.
5. D. Shepard, A two-dimensional interpolation function for irregularly spaced data, Proc. 1968 ACM nat. Conf. pp. 517-524, 1968.
6. G. F. Simmons, Introduction to topology and modern analysis, McGraw-Hill, Tokyo, 1963.