On Algorithms for Best L_2 Fits to Continuous Functions with Variable Nodes

M.J. Baines

Numerical Analysis Report 1/93

Department of Mathematics
University of Reading
P.O. Box 220
Reading, UK

The work reported here forms part of the research programme of the Oxford/Reading Institute for Computational Fluid Dynamics.

Abstract

This report gives details of a direct variational approach (with non-standard variations) used to generate algorithms to determine optimal discontinuous piecewise linear and piecewise constant L_2 fitsto a continuous function of one or two variables with adjustable nodes.Algorithms are presented which are fast and robust, and the mesh cannot tangle. An extension to higher dimensions is also given.

Contents

1 Introduction

In recent years there has been much interest in the use of irregular grids for the representation of quantities in computational modelling. This applies both to economic representation of individual features and tracking of such features as they move. Two approaches to generate such grids are through best fits with variable nodes and through equidistribution. Work on linear splines with free knots has been carried out by de Boor [4], [5], Chiu *et al.* [7], Barrow *et al.* [3] and, more recently, Loach and Wathen [10]. The equidistribution approach is described in White [12], and references therein, Kautsky & Nichols [9], Carey & Dinh [6] and Pryce [11]. A comprehensive up-to-date bibliography is given in Grosse [8].

In this report we approach the problem of finding optimal L_2 fits to continuous functions with adjustable nodes via piecewise linear discontinuous functions. Using a direct variational approach but with non-standard variations, interpreted numerically, new algorithms are devised, based on a two stage iteration process, whose limit is the required best approximation. In this way we reduce the non-linearity of the problem (eliminating it altogether for linear fits in one dimension) and obtain algorithms which are fast and robust. Using the same approach, we derive similar algorithms providing best piecewise constant L_2 fits with adjustable nodes. Fast approximate versions of the algorithms in one dimension are also given.

It is known that for continuous functions in one dimension the best piecewise linear fit amongst discontinuous functions with adjustable nodes is continuous (Chui *et al.* [7]). This result also comes out of the present analysis, except for certain cases where isolated discontinuities can occur. Thus the piecewise linear algorithm in one dimension generates piecewise linear continuous L_2 fits with adjustable nodes a.e.

In two dimensions the algorithms are based on variable triangulations of the plane, although with invariant connectivity. In this more complex case the full versions of the algorithms are less robust. We have therefore developed simplified forms which give approximately optimal piecewise linear discontinuous and piecewise constant L_2 fits with adjustable nodes on variable triangulations of a region. One of these algorithms, for piecewise constant functions, is particularly robust and successful.

The algorithms are demonstrated on various test functions. In one dimension both the full and approximate methods are fast and robust and give excellent results without any possibility of mesh tangling. In two dimensions on triangles, owing to the complexity of the problem, only the simplest algorithm is demonstrated, on functions with a single severe feature.

The plan of the report is as follows. In section 2 we obtain expeditious natural conditions in one dimension for the L_2 error between a given continuous function and a piecewise linear discontinuous function with variable nodes to have an extremum. These conditions are then used in section 3 as the basis of a new iterative algorithm designed to obtain the required best fit. The conditions also have a useful geometrical interpretation. Section 3 also contains results on two test functions. The ideas of sections 2 and 3 are repeated in section 4 for the case of piecewise constant functions with variable nodes. Approximate versions of these algorithms are then given in section 5. In section 6 natural conditions are obtained in two dimensions for the L_2 error between a given continuous function and a piecewise linear discontinuous function on a triangular grid with variable nodes to have an extremum. These are used as the basis for a two-dimensional algorithm in section 7. Once again the pattern is repeated for piecewise constant functions in section 8. This section is extended to provide a simple and robust two-dimensional algorithm. Finally, in Section 9 we give a general discussion including the extension to three dimensions and prospects for practical use. In Appendix A, the connection between such best fits and equidistribution is studied (in one dimension).

Some of this material has appeared, in embryonic form, in previous reports, but it is included again here for the sake of completeness.

2 Piecewise Linear Fits With Variable Nodes in One Dimension

Let $f(x)$ be a given C^1 function of a scalar variable x in the interval (x_o, x_{n+1}) and let $u_k(x)$ be any member of the family S_k of linear functions in the interval (x_{k-1}, x_k), where $x_o \leq x_{k-1} < x_k \leq x_{n+1}$. Then there exists a unique member $w_k^*(x)$ of S_k such that

$$\delta \int_{x_{k-1}}^{x_k} (f(x) - w_k(x))^2 dx \Big|_{w_k = w_k^*} = 0 \qquad (2.1)$$

or, equivalently,

$$\int_{x_{k-1}}^{x_k} (f(x) - w_k^*(x)) \delta w_k dx = 0 \quad \forall \delta w_k \in S_k. \qquad (2.2)$$

The function $w_k^*(x)$ is the best L_2 fit to $f(x)$ from the family S_k.

For the interval (x_o, x_{n+1}), the union of the intervals (x_{k-1}, x_k), $(k = 1, ..., n+1)$, the best L_2 fit $w^*(x)$ to $f(x)$ from the family S of piecewise linear discontinuous functions $w(x)$ with (arbitrary) jumps at $x = x_k$ $(k = 1, ..., n)$ satisfies

$$\delta \int_{x_o}^{x_{n+1}} \{f(x) - w(x)\}^2 dx \Big|_{u=u^*} =$$

$$\delta \sum_{k=1}^{n+1} \int_{x_{k-1}}^{x_k} \{f(x) - w_k(x)\}^2 dx \Big|_{w_k = w_k^*} = 0 \qquad (2.3)$$

and is also given by (2.1) or (2.2), $(k = 1, ..., n)$, since $S = \oplus S_k (k = 1, ..., n)$ and the problem decouples. The solution $w^*(x) = \cup w_k^*(x)$.

Consider now the problem of determining the best L_2 fit to $f(x)$ from the family S_D of piecewise linear discontinuous functions having arbitrary jumps at $x = x_k$ $(k = 1, ..., n)$ on a variable partition $(x_1, .., x_k, ..., x_n)$ of the fixed interval (x_o, x_{n+1}). The solution again satisfies (2.3) but, since the x_k $(k = 1, .., n)$ are to be varied as well as the w_k, the problem does not decouple in an obvious way. However, as we shall see, it is possible to regard it as two lightly coupled problems, one of which is the decoupled problem (2.2).

It is convenient at this point to introduce here a new independent variable ξ, which remains fixed, while x joins w as a dependent variable, both now depending on ξ and denoted by $\hat{x}$ and $\hat{w}$ respectively. Then, with $\hat{w}(\xi) = \hat{w}(\hat{x}_k(\xi))$, equation (2.3) becomes (reserving suffices for interval end points only)

$$\delta \sum_{k=1}^{n+1} \int_{x_{k-1}}^{x_k} \{f(\hat{x}(\xi)) - \hat{w}(\xi)\}^2 \frac{d\hat{x}}{d\xi} d\xi = 0 \tag{2.4}$$

Taking the variations of the integral in (2.4) gives

$$\int_{x_{k-1}}^{x_k} \Big\{ 2\{f(\hat{x}(\xi)) - \hat{w}(\xi)\}\{f'(\hat{x}(\xi))\delta\hat{x} - \delta\hat{w}(\xi)\} \frac{d\hat{x}}{d\xi}$$

$$+\{f(\hat{x}(\xi)) - \hat{w}(\xi)\}^2 \frac{d}{d\xi}(\delta\hat{x}) \Big\} d\xi. \tag{2.5}$$

Integrating the last term by parts leads to

$$-\int_{x_{k-1}}^{x_k} 2\Big\{ f(\hat{x}(\xi)) - \hat{w}(\xi) \Big\} \Big\{ \delta\hat{w}(\xi)\frac{d\hat{x}}{d\xi} - \frac{d\hat{w}}{d\xi}\delta\hat{x} \Big\} d\xi -$$

$$(f(\hat{x}(\xi)) - \hat{w}(\xi))_{k-1}^2 + \delta\hat{x}_{k-1} + (f(\hat{x}(\xi)) - \hat{w}(\xi))_k^2 \delta\hat{x}_k. \tag{2.6}$$

Substituting (2.6) for the last term in (2.5), collecting terms and returning to the x, w notation, (2.4) yields

$$\sum_{k=1}^{n+1} \int_{x_{k-1}}^{x_k} 2\{f(x) - w(x)\}\{\delta w - w_x \delta x\} + \sum_{j=1}^{n} \left[\{f(x) - u(x)\}^2 \right]_j \delta x_j = 0 \tag{2.7}$$

where the second summation is now over nodes j and the square bracket notation $[\ ]_j$ denotes the jump in the relevant quantity at the node j (see Fig. 10). With the constraint $\delta x = 0$ this leads back to (2.3) and (2.2) and to equations for the best piecewise linear discontinuous L_2 fit to $f(x)$ with fixed nodes.

We now discuss two kinds of variation. Choosing $\delta x = 0$ and δw to be in the space of piecewise linear discontinuous functions, (2.7) yields the conditions

$$\int_{x_{k-1}^*}^{x_k^*} \{f(x) - w_k^*(x)\}\, \phi_{ki}\, dx = 0 \quad (i = 1, 2) \tag{2.8}$$

for the best fit in element k, denoted by w_k^* and x^*, where ϕ_{k1}, ϕ_{k2} are the local linear basis functions in element k (see Fig. 1b).

As another choice, remembering that for continuity δx must lie in the space of continuous functions, we may set $\delta x = \alpha_j$ (where α_j is the standard basis function for continuous piecewise linear functions: see Fig. 1a), together with the particular constraint

$$\delta u = w_x^* \delta x, \tag{2.9}$$

in (2.7) to obtain

$$\left[\{f(x^*) - w^*(x^*)\}^2\right]_j = 0. \tag{2.10}$$

The simultaneous solution of (2.8) and (2.10) gives the required solution $w^*(x^*)$.

Using L, R for left and right values at the (variable) node j (see Fig. 1a), (2.10) may be written

$$\{f(x_j^*) - w_L^*(x_j^*)\}^2 = \{f(x_j^*) - w_R^*(x_j^*)\}^2. \tag{2.11}$$

It follows that, if $f(x_j^*) - w_L^*(x_j^*)$ and $f(x_j^*) - w_R^*(x_j^*)$ have the same sign, i.e. if w_L, w_R lie on the same side of $f(x_j^*)$ (see Fig. 2a),

$$f(x_j^*) - w_L^*(x_j^*) = f(x_j^*) - w_R^*(x_j^*)$$

$\Rightarrow$

$$w_L^*(x_j^*) = w_R^*(x_j^*), \tag{2.12}$$

irrespective of $f(x)$ (as long as it is continuous), and therefore that w^* is continuous at the new position of the node. On the other hand, if $f(x_j^*) - w_R^*(x_j^*)$ and $f(x_j^*) - w_L^*(x_j^*)$ have opposite signs, i.e. if w_L, w_R lie on the same side of $f(x_j^*)$ (see Fig. 2b),

$$f(x_j^*) - w_L^*(x_j^*) = -\{f(x_j^*) - w_R^*(x_j^*)\}$$

$\Rightarrow$

$$\frac{1}{2}\{w_L^*(x_j^*) + w_R^*(x_j^*)\} = f(x_j^*) \tag{2.13}$$

in which case w^* is discontinuous at x_j^*, its jump being bisected by bisected by $f(x_j^*)$.

Now it is known (Chui *et al.* [7], Loach & Wathen [10]) that for for continuous functions $f(x)$ the best L_2 fit amongst discontinuous piecewise linear functions with variable nodes is continuous, which clearly corresponds to (2.11). The case (2.12), with a definite discontinuity in w^* at x_j^*, therefore cannot correspond to the best least squares fit when $f(x)$ is continuous, and can correspond only to a local minimum.

3 An Algorithm for Variable Node Discontinuous Piecewise Linear Fits

An algorithm to find optimal piecewise linear L_2 fits with variable nodes can be constructed in two stages (carried out alternately until convergence is obtained), corresponding to the particular choices of variations referred to in section 2 above.
Stage (i) $\delta x_j = 0$, $(j = 1, ..., n)$,

$$\delta w = \phi_{k1} \text{ or } \phi_{k2} \quad (k = 1, ..., n+1) \tag{3.1}$$

This stage of the algorithm is governed by (2.8) and corresponds to the best L_2 fit $u(x)$ amongst the family S_D of linear functions discontinuous at <u>prescribed</u> nodes, as in (2.1),(2.2).
Stage (ii) $\delta x_j = \alpha_j$,

$$\delta w_j - w_x \delta x_j = 0 \quad (j = 1, ..., n) \tag{3.2}$$

This stage, which combines both w and x variations to give variations in w "following the motion", corresponds to finding x_j such that (2.10) holds. Geometrically, we see from (2.9) that variations of x, w are restricted to points lying on the lines of the current piecewise linear discontinuous approximation (possibly linearly extrapolated).

An algorithm is now set up, analogous to minimising a quadratic function $q(x, y)$ using two search directions d1 and d2 spanning the $x - y$ plane. Starting from some initial guess we may alternately minimise q in the directions d1 and d2. Similarly, in the present case, to find the best L_2 fit we may begin with an initial guess $\{x_j\}, \{w_j\}$; stage (i) is to find the minimum in the linear manifold specified by the variations given in (3.1) and so solve (2.8) for a new, $w(x)$ with the x_j fixed; stage (ii) is to find the minimum in the linear manifold specified by the variations given in (3.2) and so solve (2.10) approximately for new x_j by the implementation of (2.11),(2.12), as discussed below. Repetition of these stages gives a sequence which, if convergent, provides a solution of (2.4) or (2.7). As with similar problems of this type care is required as the limit may correspond only to a local minimum.

In further detail, stage (i) involves expanding $w_k(x)$ as

$$w_k(x) = w_{k1}\phi_1(x) + w_{k2}\phi_2(x) \tag{3.3}$$

(see Fig 1b) and substituting it for $w_k^*(x)$ into (2.8), yielding the equations

$$h_k \begin{bmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} w_{k1} \\ w_{k2} \end{bmatrix} = \int_{k1}^{k2} f(x) \begin{bmatrix} \phi_{k1}(x) \\ \phi_{k2}(x) \end{bmatrix} dx \tag{3.4}$$

where $h_k = x_{k2} - x_{k1}$, while stage (ii) involves taking the $w_k(x)$ which come from (3.4) and solving

$$(f(x_j) - w_{jL}(x_j))^2 - (f(x_j) - w_{jR}(x_j))^2 = 0 \tag{3.5}$$

for a new x_j (see Fig. 2).

From (2.8) we observe that $f(x) - w_k(x)$ must pass through zero at least once between x_{k-1} and x_k. Let x_{OL} and x_{OR} be the zeros closest to node j from the element either side. Then the function

$$F(x) = (f(x) - w_{jL}(x))^2 - (f(x) - w_{jR}(x))^2 \tag{3.6}$$

($c.f.$ (3.5)) has the properties

$$F(x_{0L}) < 0 \qquad F(x_{0R}) > 0, \tag{3.7}$$

(excluding the special case $w_{jL}(x_{0R}) = w_{jR}(x_{0L})$). It follows that there is at least one root of $F(x)$ between x_{0L} and x_{0R}. Choose this root (or the one nearest the old x_j if there are two) to be the new x_j. Note that if this root is chosen, all such roots lie between pairs of intersection points and mesh tangling cannot occur.

In fact, from (3.5), the new x_j must satisfy

$$(w_{jL}(x) - w_{jR}(x))(f(x_j) - \frac{1}{2}(w_{jL}(x) + w_{jR}(x))) = 0 \tag{3.8}$$

from which we either have

$$w_{jL}(x_j) = w_{jR}(x_j) \tag{3.9}$$

or

$$\frac{1}{2}(w_{jL}(x_j) + w_{jR}(x_j)) = f(x_j) \tag{3.10}$$

We shall call (3.9) the intersection construction (independent of the function $f(x)$, note) and (3.10) the averaging construction. They are represented graphically in Fig. 2.

Further information about the direction in which the nodes move may be obtained from the sign of $F(x_j)$. Solving (3.4) for w_{k1} and w_{k2} and using Simpson's rule for the integration gives

$$\begin{bmatrix} w_{k1} \\ w_{k2} \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{6}f(x_{k1}) + \frac{1}{6}f(x_{km}) + O(h_k^4) \\ \frac{1}{3}f(x_{km}) + \frac{1}{6}f(x_{k2}) + O(h_k^4) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2}{3}f(x_{k1}) + \frac{2}{3}f(x_{km}) - \frac{1}{3}f(x_{k2}) + O(h_k^4) \\ -\frac{1}{3}f(x_{k1}) + \frac{2}{3}f(x_{k2}) + \frac{2}{3}f(x_{k2}) + O(h_k^4) \end{bmatrix} \qquad (3.11)$$

where x_{km} is the mid-point of the element k. Hence

$$\left.\begin{aligned} f(x_j) - w_{jL}(x_j) &= \tfrac{1}{3}f(x_{j-1}) - \tfrac{2}{3}f(x_{j-\frac{1}{2}}) + \tfrac{1}{3}f(x_j) + O(h^4_{j-\frac{1}{2}}) \\ f(x_j) - w_{jR}(x_j) &= \tfrac{1}{3}f(x_j) - \tfrac{2}{3}f(x_{j+\frac{1}{2}}) + \tfrac{1}{3}f(x_{j+1}) + O(h^4_{j+\frac{1}{2}}) \end{aligned}\right\} \qquad (3.12)$$

and, from (3.6),

$$F(x_j) = \{\frac{1}{3}\delta^2 f_{j-\frac{1}{2}} + 0(h^4_{j-\frac{1}{2}})\}^2 - \{\frac{1}{3}\delta^2 f_{j+\frac{1}{2}} + O(h^4_{j+\frac{1}{2}})\}^2$$

$$= \frac{1}{9}(\delta^2 f_{j-\frac{1}{2}} + \delta^2 f_{j+\frac{1}{2}})(\delta^2 f_{j-\frac{1}{2}} - \delta^2 f_{j+\frac{1}{2}}) + O(\max(h^4_{j-\frac{1}{2}}, h^4_{j+\frac{1}{2}})). \qquad (3.13)$$

For sufficiently small $h_{j-\frac{1}{2}}, h_{j+\frac{1}{2}}$, therefore, the movement of node j is governed by the sign of

$$(\delta^2 f_{j-\frac{1}{2}} + \delta^2 f_{j+\frac{1}{2}})(\delta^2 f_{j-\frac{1}{2}} - \delta^2 f_{j+\frac{1}{2}}), \qquad (3.14)$$

moving left if this quantity is positive and right if it is negative.

Since in stage (ii) $w(x)$ is restricted in elements k by $\delta w = w_x \delta x$, then in elements L and R, respectively,

$$\left.\begin{aligned} w(x) - w_L &= (x - x_j)m_{k-1} \\ w(x) - w_R &= (x - x_j)m_k, \end{aligned}\right\} \qquad (3.15)$$

where $m_k = (w_x)_k$, $m_{k-1} = (w_x)_{k-1}$.

Hence, if $m_k \neq m_{k-1}$, the intersection construction (3.9) gives

$$x_j^{new} - x_j = -\left\{\frac{w_{jR} - w_{jL}}{m_k - m_{k-1}}\right\}. \qquad (3.16)$$

where x_j is the previous approximation and x_j^{new} the one currently sought. Similarly, if $m_k + m_{k-1} \neq 0$, the averaging construction (3.10) gives

$$x_j^{new} - x_j = \frac{2f(x_j^{new}) - (w_{jL} + w_{jR})}{m_k + m_{k-1}}. \tag{3.17}$$

Note that the calculation of x_j^{new} is implicit since $f(x_j^{new})$ occurs on the r.h.s.

Near to inflection points the averaging construction (3.10) may well occur (see Fig. 2b) and the fit obtained by this method will be a (discontinuous) local minimum.

For regions in which $f(x)$ is convex the new approximation to x_j is provided by the displacement (3.16), i.e. the intersections of lines in adjacent elements (see Fig. 2a), since in this case the expression $f(x) - w$ is of the same sign when approached from left or right. The fit is therefore continuous. Where $f(x)$ has an inflection point the intersection construction is replaced by the averaging construction (3.17): this occurs when the $f(x) - w$ are of opposite sign when approached from left or right, as in Fig. 2b. For these exceptional points the fit obtained by this method will be discontinuous. (One possible remedy is to change the number of points locally by one, thus breaking the symmetry, a device which seems to work well.)

In order to simplify the solution of (3.17) it is possible to make use of the outer iteration to move towards the converged x_j by using the x_j^{old} values at the previous step in the calculation of $f(x)$. In the very special case $m_{k-1} = m_k = 0$, equation (3.15) shows that x^{new} is indeterminate and there is no advantage in moving the node at all,

If $f(x)$ is convex we see from (2.12) that the result of the converged iteration (stage (i) — stage (ii) — repeated alternately) is the best <u>continuous</u> L_2 fit using piecewise linear approximation. If $f(x)$ is not convex there may possibly be isolated discontinuities in the fitted function at inflection points, where only a local minimum occurs. It is possible to replace such a discontinuous function locally by a continuous approximation, by say simply averaging the two nodal values (in which case the result is the function value itself). This is of course at the expense of abandoning the optimal fit at these isolated points. The resulting approximation may however be used as an initialisation for other algorithms completely dedicated to <u>continuous</u> best fits, see Loach & Wathen [10].

In summary the algorithm is:

1. Set up the initial grid
2. Project f elementwise into the space of piecewise linear discontinuous functions on the current grid as in (3.4) (stage (i))
3. Determine the next grid by the intersection construction (3.16) or (exceptionally) the averaging construction (3.17) (stage (ii))
4. If the new grid is too different from the previous grid, go to 2.

The algorithm, which is fast and robust, finds in appropriate cases optimal linear spline approximations with variable knots: indeed, by concentrating on piecewise linear discontinuous fits, the procedure effectively linearises the problem and avoids many of the difficulties generated by restricting the search to continuous fits at the outset.

Each step (i)+(ii) of the algorithm bears a striking resemblance to the Moving Finite Element procedure in the two step form described by Baines [2] and Baines & Wathen [1]. The similarity is pursued by Baines [2].

We show results for two examples, in Fig. 3(a),3(b).

(a)	$\tanh\{20(x-0.5)\}$	11 interior nodes
(b)	$10e^{-10x} + 20/\{1+400(x-0.7)^2\}$	9 interior nodes

In each case the fixed interval is [0,1] and the initial grid is equally spaced. Example (a) is a severe front with a single inflection. Example (b) is a test example suggested by Pryce [11].

In each example the trajectories of the nodes (i) are shown as they move towards their final positions together with the function (ii) and the fit obtained (iii). The process is taken to be converged when the ℓ_∞ norm of the nodal position updates is less than 10^{-4}. The number of iterations appears on the ordinate axis of the trajectories. In general an extra order of magnitude reduction is obtained in the L_2 error over the equispaced case.

Although the theory has been derived only for C^1 functions numerical experiments show that the algorithm also gives optimal fits to functions which are only piecewise C^1. A simple example shows that the intersection construction drives one node towards an isolated slope discontinuity (*c.f.* Fig. 2(a)), where it remains while the fits either side converge.

The algorithm also gives piecewise linear best fits to functions which have isolated discontinuities. In this case there are extra jump discontinuity terms in (2.7) arising from the variation of the integral which vanish only when a node is located at a discontinuity itself. In numerical experiments nodes move towards such a point from either side where they remain while once again the fits either side converge. This can be understood in terms of an isolated discontinuity, being a limit of a continuous steep function.

A further generalisation is to functionals of the form

$$\int F(x,u)dx \tag{3.18}$$

(c.f. (2.3)).

4 Piecewise Constant Fits With Variable Nodes in One Dimension

The approach is readily adapted for best piecewise constant fits with variable nodes. In this case the conditions for the best fit, denoted by w^*, and the grid, denoted by x^*, are

$$\int_{x_{k-1}}^{x_k} \{f(x) - w_k^*(x)\}\, \pi_k(x)\; dx = 0 \tag{4.1}$$

(c.f.(2.8)), where $\pi_k(x)$ is the characteristic function in the element k (see Fig. 1c), and

$$\left[\{f(x) - w_k^*(x)\}^2\right]_j = 0 \tag{4.2}$$

(*c.f.* (2.10)). As in section 2, equation (2.11)-(2.13), using L, R for values to the left and right of the (variable) node j, it follows from (4.2) that if w_L, w_R lie on the same side of $f(x_j^*)$

$$f(x_j^*) - w_L^*(x_j) = f(x_j^*) - w_R^*(x_j) \;\;\Rightarrow\;\; w_L^*(x_j) = w_R^*(x_j) \tag{4.3}$$

or, if w_L, w_R lie on opposite sides of $f(x_j^*)$,

$$-\left(f(x_j^*) - w_L^*(x_j)\right) = f(x_j^*) - w_R^*(x_j) \;\Rightarrow\; \frac{1}{2}(w_L^*(x_j) + w_R^*(x_j)) = f(x_j^*) \quad (4.4)$$

The latter corresponds to monotonic behaviour of f while the former exceptionally occurs near to maxima or minima. However, (4.3) gives no information about the position of x_j^* (c.f. Figs. 4a,4b). The solution is therefore the set of best constant fits in separate elements which have the averaging property (4.4).

The corresponding algorithm to find the best piecewise constant L_2 fit with variable nodes is again constructed in two stages (carried out alternately until convergence), as follows:

Stage (i) $\delta x_j = 0, \;\; (j = 1, ..., n)$,

$$\delta w = \pi_k \;\; (k = 1, 2, ..., n+1) \quad (4.5)$$

This stage of the algorithm is governed by (4.1) and corresponds to the best L_2 fit amongst the family Π_D of piecewise constant functions on a fixed grid.

stage (ii) $\delta x_j = \alpha_j$,

$$\delta w = 0 \;\; (j = 1, 2, ..., n) \quad (4.6)$$

This stage corresponds to finding x_j such that (4.2) holds, with variations of w restricted to points lying on the current piecewise constant discontinuous approximation (possibly extrapolated) in element k.

In stage (i) $w_k(x)$ is constant ($= w_k$) in each element, giving from (4.1)

$$h_k w_k = \int_{k1}^{k2} f(x)dx \quad (4.7)$$

where $h_k = x_{k1} - x_{k2}$, while stage (ii) taken the $w_k(x)$ from (4.7) and seeks x_j for which

$$(f(x_j) - w_L)^2 - (f(x_j) - w_R)^2 = 0 \quad (4.8)$$

(see Figs 4a, 4b). As in section 2 the functions $f(x) - w_L$ and $f(x) - w_R$ vanish in the intervals (x_{k-1}, x_k) and (x_{k1}, x_{k+1}) respectively. It follows that the function

$$G(x) = (f(x) - w_L) - (f(x) - w_R) \quad (4.9)$$

is negative where w intersects f in element $k-1$ and positive where w intersects f in element k (see Fig. 4). There is therefore at least one root between these points which may be chosen as the new position of x_j. Moreover, if this root is chosen, all roots (for different k) are separated by these intersections and mesh tangling cannot occur.

Further information about the direction in which the nodes move can be obtained from the sign of $G(x)$. Using the trapezium rule for the integration in (4.7) gives

$$w_k = \frac{1}{2}(f(x_{k1}) + f(x_{kL})) + O(h_k^2) \tag{4.10}$$

so that

$$\left.\begin{aligned} f(x_j) - w_{jL} &= \tfrac{1}{2}(f(x_j) - f(x_{j-1})) + O(h^2_{j-\frac{1}{2}}) \\ f(x_j) - w_{jR} &= \tfrac{1}{2}(f(x_j) - f(x_{j+1})) + O(h^2_{j+\frac{1}{2}}) \end{aligned}\right\} \tag{4.11}$$

Then, from (4.9)

$$G(x_j) = \{\frac{1}{2}(f(x_j) - f(x_{j-1})) + 0(h^2_{j-\frac{1}{2}})\}^2 -$$

$$\{\frac{1}{2}(f(x_j) - f(x_{j+1}))^2 + O(h^2_{j+\frac{1}{2}})\}^2$$

$$= \frac{1}{4}(f(x_{j-1}) - 2f(x_j) + f(x_{j+1}))(f(x_{j-1}) - f(x_{j+1})) + O(\max(h^2_{j-\frac{1}{2}}, h^2_{j+\frac{1}{2}})). \tag{4.12}$$

For sufficiently small $h_{j-\frac{1}{2}}, h_{j+\frac{1}{2}}$, therefore, the movement of node j is governed by the sign of

$$(f(x_{j-1}) - 2f(x_j) + f(x_{j+1}))(f(x_{j-1}) - f(x_{j+1})), \tag{4.13}$$

moving left if this quantity is positive and right if it is negative.

Since, in stage (ii), $w(x)$ is restricted in element k by $\delta w = 0$, then $w(x)$ is equal to the value of the current stage (i) approximation within the whole element. Hence the averaging construction (4.4) gives

$$\frac{1}{2}(w_{jL} + w_{jR}) = f(x_j) \tag{4.14}$$

c.f. (3.17), where w_{jL} and w_{jR} are the values of the current stage (i) approximation to the left and right of node j. Any standard algorithm may be used to extract x_j : here we have used bisection.

In the case of (4.3) there is no solution for x_j unless $w_{jL} = w_{jR}$. In this exceptional case any x_j is a solution and there is therefore no reason to adjust the node position at the current iteration.

In summary the algorithm is:

1. Set up an initial grid

2. Project f elementwise into the space of piecewise constant functions on the current grid as in (4.7) (stage (i))

3. Determine the new grid by the averaging construction (4.14) (stage (ii))

4. If the new grid is too different from the previous grid, go to 2.

The results are shown for the same test examples as in section 3, shown in Fig. 5(a),5(b), except that for better resolution example (b) is done with 19 interior nodes.

(a) $\tanh\{20(x-0.5)\}$	11 interior nodes
(b) $10e^{-10x} + 20/\{1+400(x-0.7)^2\}$	19 interior nodes

In both cases the interval is [0,1] and the initial grid is again equally spaced. In each example the trajectories of the nodes (i) are shown as they move towards their final positions, together with the function (ii) and the fit obtained (iii). The process is taken to have converged when the relative error in the L_2 norm of $f(x) - u(x)$ is less than 10^{-4}. The number of iterations appears on the ordinate axis of the trajectories. An order of magnitude reduction in the L_2 error over the fixed node case is obtained.

As in section 3 numerical experiments indicate that the algorithm also gives best piecewise constant fits to C^o functions which are only piecewise continuous and functions which have isolated discontinuities, with a node moving towards a discontinuity and remaining there while the rest of the fit converges.

5 Approximate Versions of the One-Dimensional Algorithms

Although the algorithm in sections 3 and 4 work perfectly satisfactorily we now describe very fast algorithms for generating approximate best fits which appear to be almost as good, for which convergence proofs can be given, and which are very useful for generalisations to higher dimensions.

These algorithms are based upon using interpolants of the function $f(x)$ at each stage of the iteration, rather than the function itself. The resulting fit is therefore not to $f(x)$ but the interpolant of $f(x)$ at the limit. The degradation is rather small, however, and the algorithms have very positive advantages.

We begin this time with the piecewise constant fits of section 4. Instead of fitting $f(x)$ we shall fit the current linear interpolant $f_I(x)$ (linear in each element) at each stage of the iteration. This means that (4.7) becomes

$$w_k = \frac{1}{2}(f(x_{k1}) + f(x_{k2})) \tag{5.1}$$

and that (4.14) becomes

$$\frac{1}{2}(w_{jL} + w_{jk}) = f_I(x_j). \tag{5.2}$$

Since $f_I(x)$ is linear in each element to the left and right of node j, it is possible to write down the solution of (5.2) for x_j (called here x_j^{new}) which (using (5.1)) is given (see Fig. 4) by

$$x_j^{new} - x_j = \frac{\frac{1}{4}(f(x_{j+1}) - 2f(x_j) + f(x_{j-1}))}{\max\left[\frac{|f(x_{j+1})-f(x_j)|}{\Delta x_{j+\frac{1}{2}}}, \frac{|f(x_j)-f(x_{j-1})|}{\Delta x_{j-\frac{1}{2}}}\right]} \tag{5.3}$$

This simple iteration replaces the two stage iteration of section 4. If it converges, the limit values satisfy

$$f(x_{j+1}^*) - f(x_j^*) = f(x_j^*) - f(x_{j-1}^*) \tag{5.4}$$

and the grid is the one that produces equi-spaced $f(x_i^*)$.

Convergence of the algorithm may be discussed via (4.13). Note that the $O(h_k^2)$ term of (4.10) is now missing so that the node j moves to the left or the right according as whether (4.13) is positive or negative. Thus, except in the vicinity of nodes where (4.13) changes sign, nodal movement is uni-directional. We may exclude the possibility of (4.13) changing sign by assigning fixed nodes to points of maxima, minima and inflection points of $f(x)$. Between these fixed points the nodes will all move in the same direction, their positions are bounded above, and they converge. The convergence result depends on these points remaining fixed.

The solution in ξ-space is simply equispaced points on the straight line

$$\hat{f}_I(\xi) = f(\xi_A)\left[\frac{\xi_B - \xi}{\xi_B - \xi_A}\right] + f(\xi_B)\left[\frac{\xi - \xi_A}{\xi_B - \xi_A}\right] \tag{5.5}$$

between any two points ξ_A and ξ_B for which the averaging construction is unique, i.e. away from maxima and minima, where

$$\hat{f}_I(\xi) = f_I(x(\xi)). \tag{5.6}$$

To solve (5.5) and (5.6) for $x(\xi)$ involves knowing in advance where the nodes $x(\xi_i)$ are. Alternatively we can solve

$$\hat{f}_I(\xi) = f(x(\xi)) \tag{5.7}$$

with (5.5), but this involves inverting the function f itself, leading again to some form of iteration.

In the piecewise linear case, instead of fitting $f(x)$ we shall fit the current quadratic interpolant in each element using the value of the function at the mid-point of the element as the third matched value. Then (3.4) gives w values (3.11) without the $O(h_k^4)$ terms, while x_j^{new} is given by (3.16), which becomes

$$x_j^{new} - x_j = \frac{\frac{1}{3}\{\delta^2 f_Q(x_{j-\frac{1}{2}}) - \delta^2 f_Q(x_{j+\frac{1}{2}})\}}{\left[\frac{f(x_{j+1})-f(x_j)}{x_{j+1}-x_j}\right] - \left[\frac{f(x_j)-f(x_{j-1})}{x_j-x_{j-1}}\right]} \tag{5.8}$$

or (3.17) in the form

$$x_j^{new} - x_j = \frac{2f_I(x_j^{new}) - (\delta^2 f_Q(x_{j-\frac{1}{2}}) + \delta^2 f_Q(x_{j+\frac{1}{2}}))}{\left[\frac{f(x_{j+1})-f(x_j)}{x_{j+1}-x_j}\right] + \left[\frac{f(x_j)-f(x_{j-1})}{x_j-x_{j-1}}\right]} \tag{5.9}$$

As in the piecewise constant case, if the iteration converges then at the limit the x_j^* satisfy a particular relationship, this time

$$\delta^2 f(x_{j-\frac{1}{2}}) - \delta^2 f(x^*_{j+\frac{1}{2}}) = 0, \tag{5.10}$$

away from inflection points.

Moreover, convergence of the algorithm is ensured as in the piecewise constant case if the $F(x_j)$ of (3.13) (without the $O(h^4)$ terms) is one-signed, which is satisfied if (3.14) is not zero. By fixing nodes at the inflection points and at points when $f^{(iv)}(x)$ vanishes, this condition is satisfied and convergence may be proved as before. Finally, we observe that for the quadratic interpolant the solution in ξ-space is equispaced points on the particular quadratic $\hat{f}_Q(\xi) = f_Q(x(\xi))$ which passes through the points $\xi_A, \frac{1}{2}(\xi_A + \xi_B), \xi_B$, between any two points ξ_A and ξ_B for which the intersection construction is unique, i.e. away from inflection points. Finding the points x_j still requires the inversion of the functions f_Q or f, however.

6 Piecewise Linear and Constant Fits in Two Dimensions

The generalisation of these techniques to two dimensions raises a number of difficulties. In principle, the same approach yields algorithms for obtaining best discontinuous fits to given continuous functions on a tessellation of the plane. The solution of the nodal position stage of the algorithm is more difficult, however, and requires additional numerical techniques. Furthermore, there is not the same simple connection in two dimensions between discontinuous linear fits and continuous ones. With these important provisos, however, we describe methods and algorithms which obtain good representation of sharp functions in two dimensions, and generalise to higher dimensions.

Let $f(x, y)$ be a given C^1 function of the two variables x and y in a domain Ω and let $w_k(x, y)$ be a member of the family S_k^2 of linear functions on a triangular subdomain Δ_k of Ω. Then there exists a unique member $w^*(x, y)$ of S_k^2 such that

$$\delta \int_{\Delta_k} \{f(x,y) - w_k(x,y)\}^2 dxdy \Big|_{w_k = w_k^*} = 0 \qquad w_k \in \ S_k \tag{6.1}$$

or, equivalently,

$$\int_{\Delta_k} \{f(x,y) - w_k^*(x,y)\} \delta w_k(x,y) \ dxdy = 0 \qquad \forall \delta w_k(x,y) \in \ S_k^2 \tag{6.2}$$

The function $w_k^*(x, y)$ is the best L_2 fit to $f(x)$ from the family S_k^2.

For the region Ω, the union of triangles Δ_k, the best L_2 fit $w_k^*(x, y)$ to $f(x, y)$ from the family S^2 of piecewise linear discontinuous functions $w_k(x, y)$ satisfies

$$\delta \int_{\Delta_k} \{f(x,y) - w^*(x,y)\}^2 dxdy =$$

$$\delta \sum_k \int_{\Delta_k} \{f(x,y) - u_k^*(x,y)\}^2 dxdy = 0 \tag{6.3}$$

and is also given by (6.1) or (6.2), since $S^2 = \oplus S_k^2$ and the problem decouples. The solution is $w^*(x,y) = \cup w_k(x,y)$.

Now consider the problem of determining the best L_2 fit to $f(x,y)$ from the family S_D^2 of piecewise linear discontinuous functions on a variable triangulation $\cup_k \Delta_k$ of the fixed domain Ω, where the internal vertices of the Δ_k are varied.

It is again convenient to introduce new independent variables ξ, η, which remain fixed, while x and y join w as dependent variables, all three now depending on ξ and η and denoted by $\hat{x}, \hat{y}$ and $\hat{w}$ respectively. Then, with $\hat{w}(\xi,\eta) = w(\hat{x}(\xi,\eta),\, \hat{y}(\xi,\eta))$, equation (6.3) becomes

$$\delta \sum_k \int_{\Delta_k} \{f(\hat{x}(\xi,\eta), \hat{y}(\hat{x}(\xi,\eta)) - \hat{w}(\hat{x}(\xi,\eta))\}^2 J \; d\xi \; d\eta = 0 \tag{6.4}$$

where $J = \frac{\partial(x,y)}{\partial(\xi,\eta)}$ is the Jacobian of the transformation.

Taking the variations of the integral in (6.4) gives

$$\int_{\Delta_k} \Big\{ 2\{f(\hat{x}(\xi,\eta)), \hat{y}(\xi,\eta)) - \hat{w}(\xi,\eta)\}$$
$$\{f_x(\hat{x}(\xi,\eta), \hat{y}(\xi,\eta))\delta\hat{x}(\xi,\eta) + f_y(\hat{x}(\xi,\eta), \hat{y}(\xi,\eta))\delta\hat{y}(\xi,\eta) - \delta\hat{w}(\xi,\eta)\}J$$
$$+ \{f(\hat{x}(\xi,\eta), \hat{y}(\xi,\eta)) - \hat{u}(\xi,\eta)\}^2 \delta J \Big\} d\xi d\eta. \tag{6.5}$$

Integrating the last term by parts leads to

$$- \int_{\Delta_k} 2\{f(\hat{x}(\xi,\eta), \hat{y}(\xi,\eta)) - \hat{w}(\xi,\eta)\}\{\delta\hat{u}J - \nabla_{\xi,\eta}\hat{w}.(\delta\hat{x}, \delta\hat{y})\} d\xi d\eta$$
$$+ \int_{\partial\Delta_k} \{f(\hat{x}(\xi,\eta), \hat{y}(\xi,\eta)) - \hat{w}(\xi,\eta)\}^2 (\delta x, \delta y).\underline{\hat{\mathbf{n}}} \; ds, \tag{6.6}$$

where $\underline{\hat{\mathbf{n}}}$ is the outward drawn normal to an element ds of the boundary $\partial\Delta_k$ of Δ_k.

Collecting terms and returning to the x, y, w notation, (6.4) yields

$$\sum_k \int_{\Delta_k} 2\{f(x,y) - w(x,y)\}\{\delta w - w_x \delta x - w_y \delta y\} dxdy \; +$$
$$\sum_k \int_{\partial\Delta_k} \{f(x,y) - w(x,y)\}^2 (\delta x, \delta y).\underline{\hat{\mathbf{n}}} \; ds = 0 \tag{6.7}$$

With the constraints $\delta x, \delta y = 0$ this leads back to (6.3) and (6.2) and to equations for the best piecewise linear discontinuous L_2 fit to $f(x,y)$ with fixed nodes.

Choosing $\delta x, \delta y = 0$ and δw to be in the space of piecewise linear discontinuous functions gives for the best discontinuous fit, denoted by w^*, x^* and y^*, the conditions

$$\int_{\Delta_k^*} \{f(x,y) - w^*(x,y)\}\phi_{k1}\ dxdy\ = 0 \quad (i = 1, 2, 3) \tag{6.8}$$

where ϕ_{k1}, ϕ_{k2}, ϕ_{k3} are local linear basis functions in the element k (see Fig. 6b). As other choices, remembering that $\delta x_j, \delta y_j$ must lie in the space of piecewise linear continuous functions, and letting α_j (see Fig. 6a) be the two-dimensional linear finite element basis function at node j, we may set (separately)

$$\left.\begin{array}{ll} & \delta x_j = \alpha_j, \delta y_j = 0, \delta w_j = w_x^* \delta x_j \\ \text{and} & \\ & \delta x_j = 0, \delta y_j = \alpha_j, \delta w_j = w_y^* \delta y_j \end{array}\right\} \tag{6.9}$$

(c.f.(2.9)) in turn in (6.7) to obtain

$$\int_{j-\text{star}} \{f(x,y) - w^*(x,y)\}^2 \alpha_j \underline{\hat{\mathbf{n}}}\ ds = 0 \tag{6.10}$$

for x_j^* and y_j^*, where $\underline{\hat{\mathbf{n}}} = (n_1, n_2)$ and "j-star" indicates the spokes, i.e. the union of the sides of the triangles, passing through the node j at x_j^*, y_j^*, (see Fig. 7).

The simultaneous solution of (6.8) & (6.10) gives the required solution $w^*(x^*, y^*)$. Note that (6.10) can be written

$$\sum_{\ell=\ell_1}^{\ell_s} \int \left[\{f(x,y) - w_{\ell A}^*(x,y)\}^2 - \{f(x,y) - w_{\ell C}^*(x,y)\}^2\right] \alpha_j \underline{\mathbf{n}}\ ds = 0 \tag{6.11}$$

where ℓ runs over the spokes ℓ_1 to ℓ_s of j-star, and $w_{\ell A}, w_{\ell C}$ refer to the values of w on the spoke looking anticlockwise and clockwise, respectively. Another useful form is

$$\begin{aligned} &\sum_{\ell=\ell_1}^{\ell_s} \int \Big\{(w_{\ell A}(x,y) - w_{\ell C}(x,y))\{(f(x,y) \\ &-\frac{1}{2}(w_{\ell A}(x,y) - w_{\ell C}(x,y))\}\Big\} \alpha_j \underline{\mathbf{n}}\ ds = 0 \end{aligned} \tag{6.12}$$

7 An Algorithm for Variable Node Discontinuous Piecewise Linear Fits on a Variable Triangulation of the Plane

An algorithm to find the best discontinuous linear L_2 fit with variable nodes is constructed in two stages (carried out repeatedly until convergence), corresponding to the choice of variations referred to in section 6 above.

Stage (i)

$$\delta x_j = 0,\ \delta y_j = 0,\ \delta w = \phi_{k1},\quad \phi_{k2}\ \text{ or }\ \phi_{k3} \tag{7.1}$$

This stage of the algorithm corresponds to the best L_2 fit amongst discontinuous piecewise linear functions on a prescribed grid, as in (6.1),(6.2), and (6.8) above.

Stage (ii), x variations

$$\delta x_j = \alpha_j,\ \delta y_j = 0,\ \delta w_j - w_x \delta x_j = 0 \quad (j = 1, 2, ..., n) \tag{7.2}$$

Stage (ii), y variations

$$\delta x_j = 0,\ \delta y_j = \alpha_j,\ \delta w_j - w_y \delta y_j = 0 \quad (j = 1, 2, ..., n) \tag{7.3}$$

Stage (ii), which combines w and x (or y) variations to give variations in w "following the motion" in the x (or y) directions, corresponds to finding x_j (or y_i) such that (6.10) (or (6.11)) holds. Geometrically, we see from (7.2) or (7.3) that variations of x, w (or (y, w)) are restricted to points lying on certain planes constructed from the stage (i) solution (possibly extrapolated) in each of the elements k surrounding j (see Fig. 7).

The problem of finding $w_k(x, y)$, belonging to S_k^2, which satisfies (6.8) is standard. Setting

$$w_k(x, y) = \sum_{i=1}^{3} w_{ki} \phi_{ki}(x, y) \tag{7.4}$$

in element k, where i ranges over the corners of Δ_k, we substitute into (6.8) and find the matrix equation

$$C_k \underline{\mathbf{w}}_k = \underline{\mathbf{b}}_k \tag{7.5}$$

where $\underline{\mathbf{w}}_k = \{w_{ki}\}, \underline{\mathbf{b}}_k = \{b_{ki}\}, b_{ki} = \int_{\Delta_k} f(x,y)\phi_{ki}\, dxdy$, and

$$C = \frac{A_k}{12} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \tag{7.6}$$

where A_k is the area of the triangular element k.

The other problems, those of finding x_j satisfying (6.10) with $\delta w_j = w_x \delta x_j$ and y_j satisfying (6.11) with $\delta w_j = w_y \delta y_j$, are more difficult non-linear problems. To make progress we shall hold the x_j in $f(x,y)$ constant in solving for the new x_j, and embed the associated iteration in the overall iteration, as in the "averaging" construction algorithm of section 3. Similarly for the y_j. This device was used in section 3 (equation (3.17)) to obtain converged solutions for x_j^*, in effect a "lagged" form of the equation being solved as the overall iteration converges.

Let $k = k_1, ..., k_e$ denote the elements surrounding the node j and let ℓ_1, ℓ_2 denote the edges of the element k emanating from node j (see Fig. 7). Then (6.12) may be written

$$\sum_{\ell=\ell_1}^{\ell_s} \int \{w_{\ell A}(x,y) - w_{\ell C}(x,y)\}$$

$$\left\{f(x,y) - \frac{1}{2}(w_{\ell A}(x,y) + w_{\ell C}(x,y))\right\}^2 \alpha_j \underline{\mathbf{n}}\, ds_\ell = 0 \tag{7.7}$$

Since $w(x,y)$ is restricted in element k by $\delta u = w_x \delta x$, $\delta y = 0$, then if w_{jk} is the value of the fit obtained from stage (i) at node j in element k, we have in element k

$$w(x,y) - w_{jk} = m_k(x - x_j) + n_k(y - y_j) \tag{7.8}$$

where $m_k = (u_x)_k$, $n_k = (u_x)_k$, to be substituted into (7.7).

This is a highly nonlinear equation, bearing in mind the dependence of the range of integration on the unknowns x_j and y_j, but if $f(x_\ell, y_\ell)$ is lagged in the iteration, it reduces to a quadratic.

We therefore introduce an iteration (to be run in tandem with the main iteration) in which we solve the first component (7.7) for $x_j^{(i+1)}$ in terms of $x_j^{(i)}$, where f, m and n are evaluated at $x_j^{(i)}$ while x_j and the range of integration are evaluated at $x_j^{(i+1)}$. This equation can then be written

$$AX^2 - BX + C = 0 \tag{7.9}$$

where

$$X = x_j^{(i+1)} - x_j \tag{7.10}$$

$$A = \sum_{\ell=\ell_1}^{\ell_s} \int (m_{\ell A}^2 - m_{\ell C}^2)\alpha_j dy_\ell \tag{7.11}$$

$$B = \sum_{\ell=\ell_1}^{\ell_s} \int \Big\{(w_{\ell A} - w_{\ell C}) + (m_{\ell A} - m_{\ell C})(x_\ell - x_j) +$$

$$(n_{\ell A} - n_{\ell C})(y_\ell - y_j)\Big\} m_k \alpha_j dy_\ell \tag{7.12}$$

$$C = \sum_{\ell=\ell_1}^{\ell_s} \int \Big[\{f(x_\ell, y_\ell) - w_{JA} - m_{\ell A}(x_\ell - x_j) - n_{\ell A}(y_\ell - y_j)\} -$$

$$\{f(x_\ell, y_\ell) - w_{jC} - m_{\ell C}(x_\ell - x_j) - n_{\ell C}(y_\ell - y_j)\}^2\Big]\alpha_j dy_\ell \tag{7.13}$$

and (provided that $B^2 > 4AC$) solved for X. The integrals in (7.13) may be evaluated by a quadrature rule. Both Gaussian quadrature and the trapezium rule have been tried. In the latter case (7.13) simplifies considerably with little degradation to the results.

Two real solutions of (7.9) may be regarded in simple situations as analogous to the "intersection" solution and "averaged" solution encountered in the 1-D case discussed in section 3, corresponding to convex or concave parts and inflection points of the function f, respectively. In the present two-dimensional case the dimensionality and the many contributions to A, B, C blur the simple 1-D interpretation but for consistency we choose the root corresponding to the least movement. If $B^2 = 4AC$ in (7.10) the roots coalesce, while if $B^2 < 4AC$ imaginary roots occur. In the latter case we go for the "nearest" real solution, which is the equal roots case.

Numerical difficulties arise when A, B and/or C become very small, which may be due to nearly plane patches in f or simply closeness to the best fit. A threshold parameter is therefore introduced which protects the roots from the resultant singularities. If $|A|, |B|$ or $|C|$ fall short of the threshold parameter, special solutions are taken. In particular, note that if $|C|$ is small we are already close to convergence.

Since the non-tangling property in one-dimension is no longer guaranteed, there may still be the possibility of nodes being carried across element boundaries, leading to triangles with negative area. In these situations a relaxation parameter is introduced which restricts each node to stay within the surrounding triangles. Even then there are rare occasions when a triangle area may go negative, in which case a local smoothing can be applied as an emergency measure, and the algorithm continued. These features greatly reduce the effectiveness of the algorithm and prompt the simplified algorithm described in section 9.

The calculation of $y^{(i+1)}$ proceeds in a similar way.

This algorithm gives an approximate optimal discontinuous linear fit on triangles. To obtain a useful continuous piecewise linear approximation we may take an average of the w_{jk} values at a given node j from each adjacent element k to give an approximate nodal value $\overline{w_j}$, or use the present approximation as a first guess in an algorithm dedicated to finding a <u>continuous</u> best fit.

In summary the algorithm is:

1. Set up the initial grid

2. Project $f(x, y)$ elementwise into the space of piecewise linear discontinuous functions on the current grid using (7.5) (stage (i))

3. Determine the next grid by solving (7.10) (and its y-direction counterpart) with a relaxation factor to prevent tangling (stage (ii))

4. If the new grid is too different from the previous grid or if the L_2 error is decreasing, go to 2.

Results are shown in Fig. 8(a-c) for three examples, each being a sharp front with a different orientation:

(a) $\tanh 20(x - \frac{1}{2})$

(b) $\tanh 20(x + y - 1)$

(c) $\tanh 20(x^2 + y^2 - \frac{1}{2})$

all on the unit square with 49 interior grid points. In each case the initial grid is uniform (Fig. 8)

Figure 8(a) shows the grid and profile for example (a) after convergence of the algorithm, while Figures 8(b) and 8(c) show the corresponding results in the case of examples (b) and (c), respectively. Note that the profiles show piecewise continuous linear plots (obtained by averaging at the nodes) whereas the true plots should be piecewise linear discontinuous.

The L_2 errors are shown in Table 1. Errors from the corresponding piecewise linear continuous function (obtained by averaging nodal values) are shown in brackets.

	Initial error	Final error	No. of steps
(a)	3.77×10^{-3}	2.37×10^{-5}	40
	(2.49×10^{-2})	(5.28×10^{-5})	
(b)	4.06×10^{-3}	5.89×10^{-6}	80
	(3.90×10^{-2})	(1.37×10^{-5})	
(c)	6.62×10^{-3}	2.43×10^{-4}	40
	(2.86×10^{-2})	(4.53×10^{-4})	

Table 1: L_2 errors for piecewise linear discontinuous best fits.

In examples (a) and (c) boundary node displacements along the boundary are set equal to the corresponding displacements on the next grid line in from the boundary. This cleans up a lot of the noise generated by the special behaviour of the boundary nodes and the resulting pollution as it spreads into the interior, giving an extra order of magnitude accuracy in this way.

8 Piecewise Constant Fits in Two Dimensions

In the case of best piecewise constant fits with adjustable nodes in two dimensions, $w_x = w_y = 0$ and (6.7) reduces to

$$\sum_k \int_{\Delta_k} 2\{f(x,y) - w(x,y)\}\delta u \, dxdy +$$

$$\sum_k \int_{\partial\Delta_k} \{f(x,y) - w(x,y)\}^2 (\delta x, \delta y).\hat{\mathbf{n}}\, ds = 0 \tag{8.1}$$

With δw as the characteristic function $\pi_k(x,y)$ on element k (Fig. 6c), and $\delta x, \delta y$ taken successively, as in section 4, to be the local "hat" function associated with node j we have that the conditions for the best piecewise constant L_2 fit to $f(x,y)$, denoted by w_k^*, x_j^* and y_j^*, are (c.f. (6.8)-(6.10))

$$\int_{\Delta_k^*} \{f(x,y) - w_k^*\}\, dxdy = 0 \tag{8.2}$$

$$\int_{j-\text{star}} \left\{ f(x,y) - \sum_{k=k_1}^{k_e} w_k^* \pi_k^*(x,y) \right\}^2 \alpha_j \underline{\mathbf{n}}\, ds = 0 \tag{8.3}$$

where j-star is as in Fig. 8, α_j is as in Fig. 7a, k runs over the elements surrounding node j and

$$w^*(x,y) = \sum_{k=k_1}^{k_e} w_k^* \pi_k^*(x,y) \tag{8.4}$$

By solving (8.2) and (8.3) simultaneously, we obtain the required fit $w^*(x,y)$

This leads to the following algorithm.

Stage (i)

$$\delta x_j = \delta y_j = 0,\ \delta w = \pi_k. \tag{8.5}$$

This stage of the algorithm corresponds to the best L_2 fit amongst piecewise constant functions on a prescribed grid ($c.f.$(8.6)).

Stage (ii), x variations

$$\delta u_j = \delta y_j = 0,\ \delta x_j = \alpha_j \tag{8.6}$$

Stage (iii), y variations

$$\delta u_j = \delta x_j = 0,\ \delta y_j = \alpha_j \tag{8.7}$$

From (8.2) stage (i) gives

$$w_k = \frac{1}{\Delta_k} \int_{\Delta_k} f(x,y)\, dxdy \tag{8.8}$$

For stage (ii) equation (8.3) may be written as ($c.f.$ (6.11))

$$\sum_{\ell=\ell_1}^{\ell_s} \int (w_{\ell A} - w_{\ell C}) \{f(x_\ell, y_\ell) - \frac{1}{2}(w_{\ell A} + w_{\ell C})\} \alpha_j\, dy_\ell = 0 \tag{8.9}$$

centred on $(x_j^{new}, y_j)(c.f.\ (7.9))$, to be solved for x_j^{new} with y_j fixed, while for stage (iii) it becomes

$$\sum_{\ell=\ell_1}^{\ell_s} \int (w_{\ell A} - w_{\ell C})\{f(x_\ell, y_\ell) - \frac{1}{2}(w_{\ell A} + w_{\ell C})\}\alpha_j \, dx_\ell = 0 \qquad (8.10)$$

centred on (x_j, y_j^{new}) to be solved for y_j^{new} with x_j fixed. (The positive sign corresponds to ℓ_1 and the negative sign to ℓ_2).

To solve (8.9), (8.10) for the new node positions x_j, y_j, respectively, we may simplify the problem by using trapezium rule quadrature and then use bisection. Again, since the non tangling property dimension is not guaranteed, a relaxation parameter must be introduced to prevent nodes crossing element boundaries. These features weaken the effectiveness of the algorithm however and prompt the approach in section 9 below.

In summary the algorithm is:

1. Set up the initial grid

2. Project $f(x, y)$ elementwise into the space of piecewise constant functions w_k in each element k as in (8.9) (stage (i))

3. Determine the new grid by solving (8.9) and (8.10) for x_j, y_j, respectively, using bisection, with a relaxation factor to prevent tangling (stage (ii))

4. If the grid is too different from the previous grid, or if the L_2 error stops decreasing, go to 2.

Results are shown in Figs. 9(a-c) for the same three examples as in section 6 on the same unit square with the same number of interior grid points. The initial grid is again uniform (Fig. 8). Figure 9(a),(b),(c) show grids and profiles for examples (a),(b),(c) after convergence of the algorithm. Note that, owing to the graphics, the Figures show piecewise continuous linear plots whereas the true plots should be piecewise constant.

The corresponding L_2 errors are shown in Table 2:

In example (a) boundary node displacements along the boundary are again set equal to the corresponding displacements on the next grid line in from the

	Initial error	Final error	No. of steps
(a)	1.8	1.55×10^{-3}	40
(b)	1.8	8.54×10^{-4}	40
(c)	1.84	2.34×10^{-3}	20

Table 2: L_2 errors for piecewise constant best fits.

boundary. Again this cleans up a lot of the noise generated by the special behaviour of the boundary nodes and the resulting pollution as it spreads into the interior.

9 Simplified Forms of the Algorithm in Two Dimensions

Now, following section 5, we develop simplified forms of the two-dimensional algorithms in sections 6-8, using the current interpolant during the iterations instead of the function itself.

We begin with the piecewise constant case of section 8. Replacing $f(x,y)$ by its linear interpolant $f_I(x,y)$, (8.9) becomes

$$w_k = \frac{1}{3}(f_{k1} + f_{k2} + f_{k3}) \tag{9.1}$$

where

$$f_{ki} = f(x_{ki}, y_{ki}), \quad i = 1,2,3 \tag{9.2}$$

and the ki are the three vertices of the triangle k.

When $f(x,y)$ is replaced by $f_I(x,y)$ in (8.9) or (8.10), the integrand is quadratic in x or y, leading to

$$\sum_{\ell=\ell_1}^{\ell_s} (w_{\ell A} - w_{\ell C})\{\frac{2}{3} f_I(x_j^{new}, y_j) + \frac{1}{3} f(x_{j\ell}, y_{j\ell}) - \frac{1}{2}(w_{\ell A} + w_{\ell C})\}(y_{j\ell} - y_j) = 0 \tag{9.3}$$

$$\sum_{\ell=\ell_1}^{\ell_s} (w_{\ell A} - w_{\ell C})\{\frac{2}{3} f_I(x_j, y_j^{new}) + \frac{1}{3} f(x_{j\ell}, y_{j\ell}) - \frac{1}{2}(w_{\ell A} + w_{\ell C})\}(x_{j\ell} - x_j) = 0 \tag{9.4}$$

where $x_{j\ell}, y_{j\ell}$ are the coordinates of the vertex on side ℓ away from the vertex j. Substituting for the $w's$ from (9.1) then gives

$$\sum_{\ell=\ell_1}^{\ell_s}(f_{\ell A}-f_{\ell C})\{f_I(x_j^{new},y_j)-\frac{1}{2}(f_{j\ell A}+f_{j\ell C})\}(y_{j\ell}-y_j)=0 \qquad (9.5)$$

$$\sum_{\ell=\ell_1}^{\ell_s}(f_{\ell A}-f_{\ell C})\{f_I(x_j,y_j^{new})-\frac{1}{2}(f_{j\ell A}+f_{j\ell C})\}(x_{j\ell}-x_j)=0. \qquad (9.6)$$

where $f_{\ell A,}f_{\ell C}$ are the values of $f(x,y)$ at the points A, C on the diagram and where

$$f_{mA}=\frac{1}{2}(f_j+f_{\ell A}), \quad f_{mC}=\frac{1}{2}(f_j+f_{\ell C}). \qquad (9.7)$$

A more general form of (9.5),(9.6) is achieved by taking variations $(\delta x,\delta y)=\alpha_j\underline{\hat{\mathbf{r}}}$ where $\underline{\hat{\mathbf{r}}}$ is a chosen direction, giving

$$\sum_{\ell=\ell_1}^{\ell_s}(f_{\ell A}-f_{\ell C})\{f_I(x_j^{new},y_j^{new})-\frac{1}{2}(f_{mA}+f_{mC})\}\Delta\ell\ \sin\theta_\ell=0 \qquad (9.8)$$

where ℓ is the length of the spoke and θ_ℓ is the angle between the spoke and the direction $\underline{\hat{\mathbf{r}}}$. The point $(x_j^{new},\, y_j^{new})$ is restricted to lie on the line through $(x_j,\, y_j)$ in the direction $\underline{\hat{\mathbf{r}}}$.

The sum in (9.8) may be regarded as a weighted sum of the curly brackets with weights

$$W_\ell=(f_{\ell A}-f_{\ell C})\Delta\ell\ \sin\theta_\ell \qquad (9.9)$$

$$f_I(x_j^{new},y_j^{new})\sum_{\ell=\ell_1}^{\ell_s}W_l=\sum_{\ell=\ell_1}^{\ell_s}\frac{1}{2}(f_{mA}+f_{mC})\,W_\ell \qquad (9.10)$$

Equation (9.10) looks very much like a generalisation of equations (4.14) or (5.2). However, the weights W_ℓ will vary and are not even positive generally. (An inspection of (9.9) shows that the W_ℓ are likely to be positive when $\underline{\hat{\mathbf{r}}}$ is in the direction of steepest f, i.e. $\underline{\nabla} f$, while if $\underline{\hat{\mathbf{r}}}$ is perpendicular to $\underline{\nabla} f$ the w_ℓ are likely to be small and of varying sign.)

The real advantage of positive weights is that, as in 1-D, we can prove a non-tangling property. For, in that case, from (9.10), $f(x_j^{new},\, y_j^{new})$ must be in the support of the values $\frac{1}{2}(f_{mA}+f_{mC})$ for all ℓ. These values are linked continuously by the spokes of j-star. Any positively averaged value will therefore intersect one of the spokes (the one with steepest f) at a point closer to (x_j, y_j) than the mid-point of the spoke, thus ensuring a displacement which cannot cause tangling.

These arguments suggest a modification to (9.10), taking $\hat{\underline{\mathbf{r}}}$ to be in the direction of the maximum slope of f along the spokes of j-star, giving $\theta = \hat{\theta}$, say, and replacing the weights W_ℓ of (9.9) in (9.10) by the positive weights

$$W_\ell^+ = |f_{\ell A} - f_{\ell C}| \, \Delta\ell \, |\sin \hat{\theta}_\ell|. \tag{9.11}$$

This modification gives the correct behaviour in the direction of steepest f but also ensures no tangling in any direction. The point (x_j^{new}, y_j^{new}) is restricted to lie on the spoke of j-star with the greatest slope.

The resulting algorithm is very simple to code and much faster and more robust than the full algorithm of section 8. Also it requires no relaxation parameter or test to see if the grid has tangled. A particularly easy version which simply takes the $W_\ell = 1$ is also viable. Graphs for the three problems of section 8 are shown in Figs. 10 (a)-(c), with the initial grid of Fig. 8, and the corresponding errors are shown in Table 3:

	Initial error	Final error	No. of steps
(a)	1.8	2.37×10^{-3}	15
(b)	1.8	2.36×10^{-3}	15
(c)	1.84	6.00×10^{-3}	15

Table 3: L_2 errors for the algorithm of section 9.

The corresponding approach for piecewise linear fits will use as interpolant a function which must be higher order than linear in any triangle but the precise choice will depend on a balance between simple quadrature and accuracy. For example, a bilinear interpolant or a full quadratic interpolant could be used, the latter being harder to integrate, the former being more subject to singularity. We shall not pursue the analysis here except to note that, by analogy with the 1-D piecewise linear case, it is the intersection construction which will dominate the iteration rather than the averaging construction above. In that spirit we can investigate an intersection of the generalisation of the 1-D piecewise linear case, namely, the intersection of the planes

$$w(x,y) = w_{jk} + m_k(x - x_j) + n_k(y - y_j) \quad \forall k \tag{9.12}$$

see (7.8)). At the "intersection" the values $w(x,y)$ are common and $x = x^*$, $y = y^*$, say. In general there is no value w_c such that

$$w_c = w_{jk} + w_k(x - x_j) + n_k(y - y_j) \quad \forall k \tag{9.13}$$

for all elements k surrounding the point j since the problem is overdetermined. A least squares solution exists, however, and is given by

$$M^T M \begin{bmatrix} w_c \\ x^* - x_j \\ y^* - y_j \end{bmatrix} = M^T \underline{\mathbf{w}}_j \tag{9.14}$$

where

$$M^T M = \begin{bmatrix} \sum 1 & -\sum m_k & \sum n_k \\ -\sum m_k & \sum m_k^2 & \sum m_k n_k \\ -\sum n_k & \sum m_k n_k & \sum n_k^2 \end{bmatrix} \tag{9.15}$$

where the sums are over surrounding k's and $\underline{w}_j = \{w_{jk}\}$.

This approach, which has much in common with the Moving Finite Element procedure, is only likely to work for fully convex $f(x,y)$, however, since if $f(x,y)$ is not convex the averaging construction may well be needed, as it is in 1-D.

10 Conclusions

We have shown that a variational approach to finding optimal L_2 fits to a continuous function among piecewise discontinuous linear or constant functions can be used to generate fast and robust algorithms for obtaining such fits. In one dimension the algorithms are simple, avoid mesh tangling and are easy to implement. If the last ounce of accuracy is not required, even simpler versions are available.

In particular, in the one-dimensional linear case the fits obtained are optimal L_2 piecewise linear continuous fits a.e.

In two dimensions the algorithms are less robust and harder to implement, needing relaxation parameters to prevent mesh tangling. Simplified versions have therefore been developed which avoid mesh tangling and hence the need for these parameters.

We demonstrate in the Appendix the strong connection between piecewise discontinuous fits in one dimension and equidistribution. The extension to three dimensions is straightforward. The main difference in the theory is that in (6.7) the two types of integral are over tetrahedra and their faces. The spokes of j-star then become the faces of the triangles which have node j as a vertex. A very simple algorithm in 3-D which avoids mesh tangling is then (9.10), the ℓ being the edges emanating from node j and with the w_ℓ taken equal to 1.

Apart from the grid generation aspects, this approach is also seen as an ingredient in an adaptive grid differential equation solver. A predicted solution on the current grid gives rise to a new grid via the fitting ideas contained in this report. The differential equation can then be re-solved with the resulting nodal movement incorporated. This is the subject of future work.

11 Acknowledgements

I am indebted to Neil Carlson for important early discussions and to Peter Sweby, Carol Reeves and Nick Birkett for programmes and programming assistance. The graphical output is from the GP2 package of Lee Buzbee.

References

[1] Baines, M.J. & Wathen, A.J. (1988). *Moving Finite Element Methods for Evolutionary Problems: Theory*, J. Comp. Phys., **79**, 245-269.

[2] Baines, M.J. (1991). *On the Relationship between the Moving Finite Element Method and the Best Fits to Functions with Adjustable Nodes*, Numerical Analysis Report 2/91, Dept. of Mathematics, University of Reading, U.K. Also in Journal for Numerical Methods in Partial Differential Equations (to appear).

[3] Barrow, D.L., Chui, C.K., Smith, P.W. and Ward, J.D., (1978). *Unicity of Best Mean Approximation by Second Order Splines with Variable Knots*, Math. Comp., **32**, 1131-1143.

[4] de Boor, C. (1973). *Good approximation by splines with variable knots*, In Spline Functions and Approximation Theory, Int. Ser. Num. Meths., **21**, Basel, Birkhauser.

[5] de Boor, C. (1974). *Good approximation by splines with variable knots II.* In: Numerical Methods for ODE's (Dundee 1973), Lecture Notes in Mathematics 263: Springer.

[6] Carey, G.F. & Dinh, H.T. (1985). *Grading Functions and Mesh Redistribution*, SIAM J. Numer. An. **22**, 1028.

[7] Chui, C.K., Smith, P.W. & Ward, J.D. (1977). *On the Smoothness of Best L_2 Approximants from Nonlinear Spline Manifolds*, Math. Comp., **31**, 17-23.

[8] Grosse, E. (1988). A Catalog of Algorithms for Approximation. In: Proceedings of the Conference on Algorithms for Approximation, Shrivenham, UK, 1988: Chapman and Hall.

[9] Kautsky, J. and Nichols, N.K. (1980). *Equidistributing meshes with constraints*, SIAM J. Sci. Stat. Comp., **1**, 449-511.

[10] Loach, P.D. & Wathen, A.J. (1991). *On the Best Least Squares Approximation of Continuous Functions using Linear Splines with Free Knots*, IMA Journal of Num. An., **11**, 393-409.

[11] Pryce, J.D. (1989). *On the Convergence of Iterated Remeshing*, IMA J. Num. An., **9**, 315-335.

[12] White, A.B. (1979). *On the selection of equidistributing meshes for two-point boundary value problems*, SIAM J. Numer. An., **16**, 473-502.

A Approximate Equidistribution Results in One Dimension

In this section, following Carey & Dinh [6], we derive asymptotic equidistribution results for the linear and constant cases in one dimension, showing the link between equidistribution and approximation by piecewise discontinuous linear and constant functions with adjustable nodes.

From (2.8) it follows that $f(x) - w^*(x)$ vanishes at at least two points in each element, s_k and t_k say. Hence $(f(x) - w^*(x))'$ vanishes at at least one point in each element, r_k say. Then, since $w^{*''} = 0$, if $f \in C^2$

$$\int_{r_k}^{x} f''(\sigma)d\sigma = \int_{r_k}^{x} (f''(\sigma) - w''(\sigma))d\sigma = f'(x) - w''(x) \tag{A.1}$$

and

$$\int_{s_k \text{ or } t_k}^{x} (f'(\lambda) - w'(\lambda))d\lambda = f(x) - w(x). \tag{A.2}$$

Hence

$$\int_{x_{k-1}}^{x_k} (f(x) - w^*(x))^2 dx = \int_{x_{k-1}}^{x_k} \left\{ \int_{s_k \text{ or } t_k}^{x} d\lambda \int_{r_k}^{\lambda} f''(\sigma)d\sigma \right\}^2 dx \tag{A.3}$$

$$\leq \int_{x_{k-1}}^{x_k} \left\{ (x_k - x_{k-1})^2 \; f''_{\text{max},k} \right\}^2 dx \tag{A.4}$$

where $f''_{\text{max},k}$ is the maximum norm of f'' in element k.

Now, if $E_1(x)$ is an equidistributing function,

$$(x_k - x_{k-1})E_1'(\theta_k) = \text{a constant}, C_1, \text{say}, \tag{A.5}$$

where $x_{k-1} < \theta_k < x_k$, and we have

$$\int_{x_o}^{x_n} (f - u)^2 dx \leq C_1^4 \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \{E_1'(\theta_k)\}^{-4} \left\{ f''_{\text{max},k} \right\}^2 dx. \tag{A.6}$$

Finally, as in Carey & Dinh [6], we approximate the right hand side of (A.6) by the integral

$$C_1^4 \int_{x_0}^{x_n} \{E_1'(x)\}^{-4} \{f''(x)\}^2 dx. \tag{A.7}$$

and minimise over functions $E_1(x)$, yielding

$$\frac{d}{dx} \left[\{E_1'(x)\}^{-5} \{f''(x)\}^2 \right] \tag{A.8}$$

$$E_1(x) \propto \int^x \{f''(\sigma)\}^{\frac{2}{5}}\, d\sigma \tag{A.9}$$

which may be regarded as the asymptotically equidistributed function.

Similarly, in the piecewise constant case, from (4.1) it follows that $f(x) - w^*(x)$ vanishes at at least one point in each element, r_k say. Then, since $w^{*'} = 0$,

$$\int_{r_k}^x f'(\sigma)d\sigma = \int_{r_k}^x (f'(\sigma) - w^{*'}(\sigma))d\sigma = f(x) - u_k^*(x) \tag{A.10}$$

Hence

$$\int_{x_{k-1}}^{x_k} (f(x) - w_k^*)^2 dx = \int_{x_{k-1}}^{x_k} \left\{ \int_{r_k}^x f'(\sigma)d\sigma \right\}^2 dx \tag{A.11}$$

$$\leq \int_{x_{k-1}}^{x_k} \left\{ (x_k - x_{k-1}) f'_{\text{max},k} \right\}^2 dx \tag{A.12}$$

where $f'_{\text{max},k}$ is the maximum norm of f' in element k.

Now, if $E_o(x)$ is the equidistributing function,

$$(x_k - x_{k-1})E'_o(\theta_k) = \text{a constant}, C_o, \text{say}, \tag{A.13}$$

where $x_{k-1} < \theta_k < x_k$, and we have from (A.12)

$$\int_{x_o}^{x_n} (f - w_k)^2 dx \leq C_0^2 \sum_{k=1}^n \int_{x_{k-1}}^{x_k} \{E'_0(\theta_k)\}^{-2} \left\{ f'_{\text{max},k} \right\}^2 dx. \tag{A.14}$$

Finally, as before, we approximate the right hand side of (A.14) by the integral

$$C_0^2 \int_{x_0}^{x_n} \{E'_0(x)\}^{-2} \{f'(x)\}^2 dx. \tag{A.15}$$

and minimise over functions $E_o(x)$, yielding

$$\frac{d}{dx}\left[\{E'_0(x)\}^{-3} \{f'(x)\}^2 \right] \tag{A.16}$$

or

$$E_0(x) \propto \int^x \{f'(\sigma)\}^{\frac{2}{3}}\, d\sigma \tag{A.17}$$

which may be regarded as the asymptotically equidistributed function.

These results are approximately borne out by the results obtained in sections 3 and 4, which therefore correspond to approximate equidistribution of the functions (A.9) and (A.17), respectively.

Figure Captions

1. Basis Functions in One Dimension

2. Linear Fits to (a) Convex and (b) Non-convex Functions

3. Results for Piecewise Linear Fits in One Dimensions
   (i) Trajectories (ii) Function (iii) Fit

4. Constant Fits to (a) Monotonic and (b) Non-monotonic Functions

5. Results for Piecewise Constant Fits in One Dimension
   (i) Trajectory (ii) Function (iii) Fit

6. Basis Functions in Two Dimensions

7. Node Connections in Two Dimensions

8. Results for Piecewise Linear Fits in Two Dimensions

9. Results for Piecewise Constant Fits in Two Dimensions

10. Results for the simplified algorithm of section 9 for Piecewise Constant Fits in Two Dimensions.