Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 202 (2005) 488–506
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

The embedded finite difference method for the Poisson equation in a domain with an irregular boundary and Dirichlet boundary conditions

Z. Jomaa, C. Macaskill *

School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia

Received 11 December 2003; received in revised form 17 June 2004; accepted 14 July 2004
Available online 26 August 2004

Abstract

The Poisson equation subject to Dirichlet boundary conditions on an irregular domain can be treated by embedding the region in a rectangular domain and solving using finite differences over the rectangle. The crucial issue is the discretization of the boundaries of the irregular domain. In the past, both linear and quadratic boundary treatments have been used and error bounds have been derived in both cases, showing that the linear case gives uniform second-order accuracy, whereas the quadratic case gives third-order accuracy at the boundaries and second-order accuracy internally. Thus, it has been recommended that the linear boundary treatment be used, as it is simpler, gives rise to a symmetric matrix formulation and has uniform accuracy. The present work shows that this argument is inadequate, because the coefficients of the error terms also play an important role. We demonstrate this in the 1-D case by determining explicit expressions for the error for both the linear and quadratic boundary treatments. It is shown that for the linear case the coefficient of error is in general large enough to dominate the calculation and that therefore it is necessary to use a quadratic boundary treatment in order to obtain errors comparable with those obtained for a regular domain. We go on to show that the 1-D expressions for error can be used to approximate the boundary error for 2-D problems, and that for the linear treatment, the boundary error again dominates.

1. Introduction

In this paper, we consider the solution of the Poisson equation on an irregular 2-D domain, subject to Dirichlet boundary conditions. The approach used here goes back to the early works of Collatz [3] and of

* Corresponding author.
E-mail address: c.macaskill@maths.usyd.edu.au (C. Macaskill).

0021-9991/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2004.07.011

Shortley and Weller [21], in that the region of interest is embedded in a rectangular domain and the boundary interpolation schemes used in the two coordinate directions are independent. At internal grid points the Poisson equation is discretized with the standard second-order accurate five-point finite difference discretization, but special treatment is required at the edges. This gives rise to a pentadiagonal matrix system. In [3], a first-order approximation is used for the estimation of derivatives at the boundary in both the x- and y-directions. For the constant coefficient Poisson equation discussed here, this scheme is equivalent to that used in [7]. However, the work of [7] can also be viewed as a development of that of [11], where the authors describe a first-order accurate symmetric discretization of the variable-coefficient Poisson equation in the presence of an irregular interface across which the variable coefficient, the solution and the derivative of the solution may have jumps.

We note that there are many other approaches to this problem in the literature. The immersed boundary method [20] uses a δ-function on the domain boundary to enforce a no flow boundary condition; see [19] for details. A related approach called the immersed interface method is a second-order numerical method designed to preserve the jump condition at the interface [10]. An alternative approach, using boundary integral techniques, has been explored in a sequence of papers [14–17].

The approach of [3,7] leads to a method that is uniformly second-order accurate, and in addition gives rise to a symmetric matrix for the representation of the second derivative operator, as only the diagonal entries are modified when the boundary points are taken into account. Similarly [8] uses a linear fit at the boundary but for more general orthogonal meshes. In [7], there is an explicit formula for the quadratic boundary treatment of the 1-D case that is the same as that used in the present work, but it is not discussed further there because a non-symmetric matrix is involved. (This quadratic formulation in fact goes back to [21] and is also mentioned in [4].) This would appear to be the end of the story, but as shown here, using the linear method the coefficient of error at the boundaries is significantly greater than at internal grid points, so that the boundary error dominates the calculation. One is thus led to a higher-order discretization at the boundaries, as in [9], who used a local area fit to obtain the appropriate second-order accurate treatment of derivatives at the boundary (i.e., using a quadratic fit). However, we employ the simpler technique of applying the boundary treatment over x and y separately as in the original quadratic scheme of [21]. In practice, the Shortley–Weller approach gives the same order error as in [9], but with a somewhat smaller coefficient of error, with the added advantage that the method (in principle) extends immediately to three dimensions, as mentioned in [21]. We note here that [22] also use quadratic fitting at the boundaries in their treatment of moving interface problems, but their method is closer to that of [9].

There is an extensive body of work giving bounds for the error involved in either the Collatz (linear) or Shortley–Weller (quadratic) boundary fitting techniques. In the first work on this problem Gerschgorin [6] used a constant fit, setting the solution value at interior grid points next to the boundary equal to the value of the nearest boundary point. Also, error bounds for difference approximations of elliptic problems were first derived by him in this same paper. Using a discrete analogue of the maximum principle for Laplace's equation, he shows that his method has errors of $O(h)$, where h is the grid spacing. In 1933, Collatz [3] improved Gerschgorin's approximation by using a linear fit along the boundary. His paper uses Gerschgorin estimation to show that the new method is $O(h^2)$. The estimates of [6,3] assume the knowledge of bounds for certain higher derivatives of the solution of the Dirichlet problem. In [21], it is shown that use of a quadratic fit at the boundary gives errors of $O(h^2)$, with the internal error $O(h^2)$ and the boundary error $O(h^3)$, so that the internal error dominates. All these cases were confirmed in [23], while in [2] similar results were obtained using Green's third identity. More generally, in [18] it is shown that a method that uses a polynomial of degree less than or equal to 6 to perform the fit along the boundary gives $O(h^2)$ convergence, again because of the internal error. A useful short historical review of this early work is given in [1]. More recently, the Shortley and Weller approximation was revisited in [13,5], these authors concentrating on the $O(h^3)$ accuracy near the boundary previously noted by others.

These asymptotic results can be understood in elementary terms. If we consider first the truncation error at the boundary for the Shortley–Weller case where a quadratic fit is used for local slopes (first derivatives), we expect that second derivatives are only accurate to first order. However, integrating twice produces a local third-order error at the boundary, as was noted in [2]. Similarly, as internal second derivatives are locally second-order accurate we initially expect fourth-order accuracy, after integrating twice. However, as will be shown explicitly for the 1-D version of the problem, the sum of N^2 such terms returns internal second-order accuracy, in agreement with the asymptotic results discussed above. Thus, we apparently have a scheme of inconsistent order, with cubic accuracy at the boundaries and quadratic accuracy internally, whereas for the linear Collatz case, the error is uniformly $O(h^2)$. In the absence of explicit coefficients for the error terms, it is not possible to decide whether the linear or quadratic method is preferable. In particular, many authors suggest use of the Collatz (linear) technique on the basis that it is more efficient and gives the same order of error as the quadratic Shortley–Weller approach. However, if one uses the Collatz linear approximation for derivatives at the boundary, then boundary error dominates the results, both for 1-D and 2-D examples. In practice, this means that in 1-D the linear boundary treatment may need as many as three times as many grid points as the quadratic boundary treatment for the same maximum absolute error. This result scales with the dimensionality of the problem, so that in 2-D as many as 10 times as many grid points may be required using the linear treatment at the boundary. The downside of the quadratic treatment of boundary points is that the symmetric matrix representation of the second derivative operator is lost, with boundary corrections appearing in all five diagonals of the matrix.

In the remainder of this paper, we describe the algorithm, provide analysis of the error involved and give some numerical examples. In particular, we first describe the discretization in 1-D and give exact explicit expressions for the coefficients of error in the linear Collatz and quadratic Shortley–Weller methods, confirming and extending the known general results, using techniques based on those of [9], but extending these to treat both the linear and quadratic cases and clearly demonstrating the source of the various error terms. In addition, we explicitly allow for general values of α_L and α_R, whereas $\alpha_L = \alpha_R = 1/2$ in [9]. We then discuss the 2-D algorithm and finally, we show how the 1-D error analysis for the linear case gives an approximate description of the errors involved in applying the method in 2-D and illustrate this with some numerical results.

2. Mathematical formulation

The method used here follows that of [3]. We consider first, in detail, the 1-D analogue of the Poisson equation. There are two reasons for this. The first is that the 2-D Poisson equation (and in principle the 3-D equation) is treated by applying the 1-D formulation of the boundary conditions separately in each coordinate direction. The second is that for the 1-D problem an exact error analysis can be carried out for both linear and quadratic treatment of the boundaries, which allows a complete understanding of the rate of convergence and demonstrates why the linear treatment is in general inadequate. Because the 2-D problem is solved by applying the 1-D method in each coordinate direction separately, the 1-D error analysis can also be applied (approximately but quantitatively) to the 2-D problem. As we show, the dominant role of the boundary error when a linear boundary treatment is used turns out to be completely analogous to the 1-D case.

2.1. One-dimensional case

We consider the domain $x \in [a,b]$ with the 1-D Poisson equation

$$\frac{d^2\psi}{dx^2} = f(x) \tag{1}$$

assumed to hold on the interior interval $x \in [x_L, x_R]$. A uniform grid is taken over $[a,b]$. Dirichlet boundary conditions are assumed given at the two boundary points $x = x_L, x_R$, which typically are not grid points. Outside the interior interval we set $\psi = 0$, so that in general there is a discontinuity at each of x_L and x_R: with zero boundary conditions this will reduce to a jump in slope. Then we label the points between the jumps so that $a = x_0 < x_L < x_1 < x_2 < \cdots < x_{N-2} < x_{N-1} < x_R < b = x_N$, with $x_L - x_0 = \alpha_L \Delta x$ and $x_N - x_R = \alpha_R \Delta x$.

The discretization of Eq. (1) gives rise to a tridiagonal matrix equation for the unknown ψ on the interior grid-points; for the exterior grid-points we set $\psi = 0$. At each interior grid-point $x = x_k$, $k = 2, \ldots, N-2$, (1) is discretized using the standard centered finite difference approximation:

$$\left[\left(\frac{\psi_{k+1} - \psi_k}{\Delta x}\right) - \left(\frac{\psi_k - \psi_{k-1}}{\Delta x}\right)\right] \Big/ \Delta x = f_k. \tag{2}$$

To complete the formulation, we require the discretization of (1) at $x = x_1$ and $x = x_{N-1}$, where the second derivative approximation must be modified to account for the boundary jumps. Consider the left-hand end $x = x_1$ (the right-hand end treatment is analogous). Here, we can use either a linear or quadratic treatment to approximate $d\psi/dx$ at $x = x_{1/2}$. Fitting a linear polynomial through the values $\psi_L = \psi(x_L)$ and ψ_1 and evaluating the corresponding slope at $x = x_{1/2}$ gives the discretization of Eq. (1) at $x = x_1$ as

$$\left[\frac{\psi_2 - \psi_1}{\Delta x} - \frac{\psi_1 - \psi_L}{(1 - \alpha_L)\Delta x}\right] \Big/ \Delta x = f_1. \tag{3}$$

The fact that only one off-diagonal element of the matrix is modified at each boundary means that it is straightforward to rewrite the governing system in symmetric form, as has been discussed elsewhere [7].

Alternatively, fitting a quadratic polynomial through the values ψ_L, ψ_1 and ψ_2 and using this to estimate $d\psi/dx$ at $x = x_{1/2}$ gives the discretization of (1) at $x = x_1$ as

$$\left\{\frac{\psi_2 - \psi_1}{\Delta x} - \frac{1}{\Delta x}\left[-\frac{2}{(1 - \alpha_L)(2 - \alpha_L)}\psi_L + \frac{1 + \alpha_L}{1 - \alpha_L}\psi_1 - \frac{\alpha_L}{2 - \alpha_L}\psi_2\right]\right\} \Big/ \Delta x = f_1. \tag{4}$$

This appears to be intrinsically non-symmetric.

2.2. Error estimation for the 1-D case

2.2.1. Linear boundary treatment

For the 1-D problem, it is possible to obtain explicit expressions for the error $\xi = \psi - \psi^e$, where ψ is the numerical solution (a function of the number of grid points and the nature of the discretization) and ψ^e is the exact solution of (1). The approach used here follows that outlined in [9], but not explicitly carried right through there. We present results for both the linear and quadratic boundary treatments. However, for the linear case, we find that further analytic approximations can be made, allowing us to identify separately the errors due to the boundary treatment and the internal discretization.

We define the second derivative operator acting on the required error ξ as

$$(L\xi)_{i+1} = \frac{H_{i+3/2} - H_{i+1/2}}{\Delta x} = \tau_{i+1}, \tag{5}$$

where H is the first derivative of ξ. The truncation error τ is defined as

$$\tau_i = f_i - (L\psi^e)_i.$$

By expanding around $x = x_1$, we find the truncation error at the left-hand boundary is

$$\tau_1 = \frac{1}{2}\alpha_L \psi_1'' - \frac{\alpha_L}{2}\left(\frac{3}{4} - \frac{\alpha_L}{2} + \frac{\alpha_L^2}{3}\right)\Delta x \psi_1''' + \left[-\frac{1}{12} + \alpha_L\left(\frac{1}{8} - \frac{\alpha_L}{8} + \frac{\alpha_L^2}{12}\right)\right](\Delta x)^2 \psi_1^{(4)} + O((\Delta x)^3), \tag{6}$$

where $\psi_1' = \psi'(x_1)$ and so on.

Similarly, for the right-hand boundary

$$\tau_{N-1} = \frac{1}{2}\alpha_R \psi''_{N-1} - \frac{\alpha_R}{2}\left(\frac{3}{4} - \frac{\alpha_R}{2} + \frac{\alpha_R^2}{3}\right)\Delta x \psi'''_{N-1} + \left[-\frac{1}{12} + \alpha_R\left(\frac{1}{8} - \frac{\alpha_R}{8} + \frac{\alpha_R^2}{12}\right)\right](\Delta x)^2 \psi^{(4)}_{N-1} + O((\Delta x)^3). \tag{7}$$

At internal grid points, we have the simple expression

$$\tau_k = -\frac{(\Delta x)^2}{12}\psi^{(4)}_k + O((\Delta x)^4), \quad 2 \leqslant k \leqslant N-2. \tag{8}$$

The error $\xi = \psi - \psi^e$ satisfies the following system of equations:

$$L\xi = \tau, \quad \xi_L = \xi_R = 0.$$

Rearranging this system of equations, we obtain:

$$H_{i+1/2} - H_{i-1/2} = \tau_i \Delta x, \quad i = 1, \ldots, N-1. \tag{9}$$

Hence

$$H_{i+1/2} = H_{N-1/2} - \Delta x \sum_{k=i+1}^{N-1} \tau_k, \quad i = 0, \ldots, N-2. \tag{10}$$

From the linear boundary treatment of $\psi'_{1/2}$, used in (3) we find

$$\xi_1 = (1 - \alpha_L)\Delta x H_{1/2} \tag{11}$$

and hence, on taking the sum of the $N-1$ equations in (9) we find

$$\xi_1 = (1 - \alpha_L)(\Delta x)^2 \left[\frac{H_{N-1/2}}{\Delta x} - \sum_{k=1}^{N-1} \tau_k\right]. \tag{12}$$

The remaining error terms can be determined using the recurrence

$$\xi_k - \xi_{k-1} = \Delta x H_{k-1/2}, \quad k = 2, \ldots, N-1. \tag{13}$$

The sum of the $N-2$ equations in (13) gives

$$\xi_k = \xi_1 + \Delta x \sum_{m=1}^{k-1} H_{m+1/2} \tag{14}$$

and then use of (9) gives

$$\xi_k = (\Delta x)^2 \left[(k - \alpha_L)\frac{H_{N-1/2}}{\Delta x} - (1 - \alpha_L)\sum_{k=1}^{N-1}\tau_k - \sum_{m=1}^{k-1}\sum_{j=m+1}^{N-1}\tau_j\right]. \tag{15}$$

Finally, using (15) with $k = N-1$ and the linear approximation for ψ'_{N-1} gives

$$H_{N-1/2} = \frac{\Delta x}{N - \alpha_L - \alpha_R}\sum_{j=1}^{N-1}(j - \alpha_L)\tau_j \tag{16}$$

and hence we have a closed form for the errors ξ_k in terms of the truncation errors τ_i.

The double summation in (15) is unwieldy and can be replaced by the following simpler form:

$$\xi_k = (\Delta x)^2 \left[\left(\frac{k - \alpha_L}{N - \alpha_L - \alpha_R} - 1\right)\sum_{j=1}^{N-1}(j - \alpha_L)\tau_j - \sum_{j=k+1}^{N-1}(k - j)\tau_j\right]. \tag{17}$$

Here the total error at the end-points ξ_1 and ξ_{N-1} is $O((\Delta x)^2)$. For $2 \leqslant k \leqslant N-2$ the error at internal grid-points, ξ_k is also $O((\Delta x)^2)$, so that we have apparently uniform error. (Note that a sum like $\sum j\tau_j$ is $O(1)$, corresponding to N^2 terms of $O((\Delta x)^2)$.) In practice the end-point error dominates, since the coefficients for ξ_1 and ξ_{N-1} are usually numerically significantly greater than the coefficients of error at the internal points. We investigate this in more detail in the next sub-section.

2.2.2. Analysis of the linear error expressions

We show in this section that simplified expressions can be obtained for the end-point error when the linear boundary treatment as in (3) is used. In addition, an integral approximation can be found for the internal error.

For convenience, we consider the special case where the right-hand boundary coincides with a grid-point, i.e., $\alpha_R = 0$, so that the dominant boundary error comes from the left-hand end. In this case, we have from (6)

$$\tau_1 = \frac{\alpha_L}{2}\psi_1'' + O(\Delta x) \tag{18}$$

and

$$\tau_i = -\frac{(\Delta x)^2}{12}\psi_i^{(4)}, \qquad i = 2, \ldots, N-1. \tag{19}$$

Hence

$$\xi_k = (\Delta x)^2\left[\left(\frac{1}{1-\alpha_L/N}\right)\left(\frac{k}{N}-1\right)\sum_{j=1}^{N-1}(j-\alpha_L)\tau_j - \sum_{j=k+1}^{N-1}(k-j)\tau_j\right]. \tag{20}$$

To identify the error due to the jump at the left-hand end-point, we compare Eq. (20) (where $\alpha_L \neq 0$) with the following simplified form with $\alpha_L = 0$:

$$\xi_k(0) = (\Delta x)^2\left[\left(\frac{k}{N}-1\right)\sum_{j=1}^{N-1} j\tau_j - \sum_{j=k+1}^{N-1}(k-j)\tau_j\right]. \tag{21}$$

Taking the difference between Eqs. (20) and (21), and retaining only the highest order terms, gives the following approximation for the left-hand end-point contribution to the error:

$$\xi_k - \xi_k(0) = \left(\frac{k}{N}-1\right)(1-\alpha_L)\tau_1(\Delta x)^2 = \frac{(k-N)\alpha_L(1-\alpha_L)}{2N}\psi_1''(\Delta x)^2. \tag{22}$$

Taking the derivative of Eq. (22) with respect to α_L shows that the maximum end-point error occurs when $\alpha_L = 0.5$, which is not surprising. Furthermore, the end-point contribution vanishes when the jump coincides with a grid-point, i.e., when $\alpha_L = 0,1$. In addition, the expression (22) is linear in x, with absolute maximum occurring at the left-hand end, and diminishing to zero at the right-hand end.

One can derive a similar expression for the error at the right-hand end:

$$\xi_k - \xi_k(0) = -\frac{k}{N}(1-\alpha_R)\tau_{N-1}(\Delta x)^2 = -\frac{k\alpha_R(1-\alpha_R)}{2N}\psi_{N-1}''(\Delta x)^2. \tag{23}$$

There does not seem to be any simple way to carry out an analogous calculation for quadratic treatment of the boundaries, because the end-point error is no longer dominant and hence does not appear at leading order.

From Eqs. (22) and (23), since the maximum end-point error occurs when $\alpha = 0.5$, we can give an approximate upper bound for the magnitude of the end-point error:

$$\xi_{\text{end}}^{\max} = \tfrac{1}{2}\max(|\,\psi''(x_{\mathrm{L}})\,|, |\,\psi''(x_{\mathrm{R}})\,|)(\Delta x)^2. \tag{24}$$

We can also obtain a simpler approximation for the error in the absence of end-point jumps in Eq. (21). This is obtained by interpreting the summations appearing in (21) as Riemann sums and then replacing them with integrals:

$$\xi_k(0) = \frac{1}{3N^2}\left[\int_{x_{\mathrm{L}}}^{x_k} x\psi^{(4)}(x)\,\mathrm{d}x - \frac{x_k - x_{\mathrm{L}}}{x_{\mathrm{R}} - x_{\mathrm{L}}}\int_{x_{\mathrm{L}}}^{x_{\mathrm{R}}} x\psi^{(4)}(x)\,\mathrm{d}x + x_{\mathrm{L}}\frac{x_k - x_{\mathrm{L}}}{x_{\mathrm{R}} - x_{\mathrm{L}}}\int_{x_{\mathrm{L}}}^{x_{\mathrm{R}}} \psi^{(4)}(x)\,\mathrm{d}x \right.$$
$$\left. - x_{\mathrm{L}}\int_{x_{\mathrm{L}}}^{x_k} \psi^{(4)}(x)\,\mathrm{d}x + (x_k - x_{\mathrm{L}})\int_{x_k}^{x_{\mathrm{R}}} \psi^{(4)}(x)\,\mathrm{d}x\right]. \tag{25}$$

This expression, which represents the *total* error in the absence of boundary jumps, can also be viewed as a close approximation to the total error obtained even in the presence of boundary jumps when a quadratic boundary fit is employed, because in that case the boundary error is $\mathrm{O}((\Delta x)^3)$ and makes a negligible contribution.

An example makes it clear how (24) and (25) can be used. If $-x_{\mathrm{L}} = x_{\mathrm{R}} = \beta$ and $\psi = \mathrm{e}^{mx}$ for some integer m, then

$$\xi_k(0) = \frac{m^2}{6\beta N^2}\left[(x_k + \beta)\mathrm{e}^{m\beta} + (\beta - x_k)\mathrm{e}^{-m\beta} - 2\beta\mathrm{e}^{mx_k}\right]. \tag{26}$$

The maximum value of this expression occurs near $x_k = 0$ (more precisely, at $x = \log(\sinh m\beta/m\beta)/m$). Therefore, the ratio defined by

$$r = \left|\frac{\xi_{\text{end}}^{\max}}{\xi_k^{\max}(0)}\right| \tag{27}$$

takes the form

$$r = 3\left|\frac{\mathrm{e}^{m\beta}}{\mathrm{e}^{m\beta} + \mathrm{e}^{-m\beta} - 2\beta}\right| \tag{28}$$

for this case (if we assume $x_k = 0$). The expression (28) gives a ratio of maximum end-point error to maximum internal error that is relatively large, e.g., $r \approx 7.5$ when $m = 1$ and $\beta = 1$, $r \approx 3.9$ when $m = 1$ and $\beta = 0.5$. Other forms for ψ give less dramatic results, so that for example $\psi = \cos x$ with $\beta = 1$ gives $r \approx 1.8$ and $\psi = \sin 3x$ with $\beta = 0.7$ gives $r \approx 2.9$. (Clearly, with sinusoidal solutions ψ it is possible to choose examples such that expression (24) for the end-point error is identically zero.)

2.3. Quadratic boundary treatment

The analysis leading to the form (17) for the errors arising from a linear boundary treatment can be repeated for the quadratic boundary fitting described in Eq. (4). We merely quote the results here.

The truncation errors are:

$$\tau_1 = -\frac{\Delta x}{3}\alpha_{\mathrm{L}}\psi_1''' - \frac{(\Delta x)^2}{12}(1 - 2\alpha_{\mathrm{L}})\psi_1^{(4)}, \tag{29}$$

$$\tau_{N-1} = \frac{\Delta x}{3}\alpha_{\mathrm{R}}\psi_{N-1}''' - \frac{(\Delta x)^2}{12}(1 - 2\alpha_{\mathrm{R}})\psi_{N-1}^{(4)} \tag{30}$$

and

$$\tau_k = -\frac{(\Delta x)^2}{12}\psi_k^{(4)}, \quad 2 \leqslant k \leqslant N - 2. \tag{31}$$

Following the approach outlined above for the linear boundary treatment, we obtain:

$$\xi_k = (\Delta x)^2 \left[[(k-\alpha_L)\frac{H_{N-1/2}}{\Delta x} - \frac{1}{2}(1-\alpha_L)(2-\alpha_L)\tau_1 + \sum_{j=1}^{N-1}(\alpha_L - j)\tau_j - \sum_{j=k+1}^{N-1}(k-j)\tau_j \right], \tag{32}$$

where

$$H_{N-1/2} = \Delta x \left[(\alpha_L - 1)(2-\alpha_L)\tau_1 + [2(1-N+\alpha_L) + \alpha_R(\alpha_R - 1)]\tau_{N-1} + 2\sum_{j=2}^{N-2}(\alpha_L - j)\tau_j \right] \times [2N - 2\alpha_R - 2\alpha_L]^{-1}. \tag{33}$$

Here, $H_{N-1/2}$ is $O((\Delta x)^2)$, so that the total error at the end-points ξ_1 and ξ_{N-1} is $O((\Delta x)^3)$. However, for $2 \leqslant k \leqslant N-2$ the error at internal grid-points ξ_k is $O((\Delta x)^2)$. These orders are the same as those found in [9], who used a slightly different quadratic boundary treatment, but as our numerical results will show, the present method actually gives end-point errors with smaller coefficients, so that even for moderate N the internal error dominates, and the convergence is uniformly $O((\Delta x)^2)$.

2.4. Overview of the 1-D error analysis

In Table 1, we give an overview of the above analysis. Column (a) of the table refers to the standard problem, where the boundaries coincide with grid points, column (b) is the embedded boundary problem, with a linear treatment at the boundaries and (c) is the embedded boundary problem with quadratic treatment of the boundaries. Internal truncations errors τ_i are $O((\Delta x)^2)$ in each case. However, the boundary truncation errors τ_1 and τ_{N-1} move from O(1) for the linear boundary treatment to $O(\Delta x)$ for the quadratic boundary treatment and finally $O((\Delta x)^2)$ when there are no boundary jumps. Inversion of the Poisson equation corresponds to integrating the truncation error twice, so that the boundary errors ξ_1 and ξ_{N-1} then range from $O((\Delta x)^2)$ for the linear boundary treatment through to $O((\Delta x)^4)$ for the no-jump problem. By contrast internal errors in all three cases are $O((\Delta x)^2)$, because of the double summation (i.e. approximately N^2 terms) each of $O((\Delta x)^4)$. Therefore, the linear boundary treatment gives rise to boundary errors that are of the same order as the internal error, but the large coefficients of these errors in most cases means that it is necessary to move to the quadratic treatment at the boundaries.

2.5. Two-dimensional case

Consider the two-dimensional Poisson equation

$$\nabla^2 \psi = f(x,y), \tag{34}$$

Table 1
The order of the error for the three 1-D cases: (a) no boundary jumps; (b) linear treatment of boundary jumps and (c) quadratic treatment of boundary jumps

(a) No jumps	(b) Linear	(c) Quadratic
$\tau_i = O((\Delta x)^2), \quad i = 1,\ldots,N-1$	$\tau_1 = O(1)$ $\tau_{N-1} = O(1)$ $\tau_i = O((\Delta x)^2), \quad i = 2,\ldots,N-2$	$\tau_1 = O(\Delta x)$ $\tau_{N-1} = O(\Delta x)$ $\tau_i = O((\Delta x)^2), \quad i = 2,\ldots,N-2$
$L\xi = \tau$ $\xi_1 = O((\Delta x)^2)\ \tau_1 = O((\Delta x)^4)$ $\xi_{N-1} = O((\Delta x)^4)$ $\xi_i = O((\Delta x)^2)\sum\sum \tau_i = O((\Delta x)^2)$, $i = 2,\ldots,N-2$	$L\xi = \tau$ $\xi_1 = O((\Delta x)^2)\ \tau_1 = O((\Delta x)^2)$ $\xi_{N-1} = O((\Delta x)^2)$ $\xi_i = O((\Delta x)^2)\sum\sum \tau_i = O((\Delta x)^2)$, $i = 2,\ldots,N-2$	$L\xi = \tau$ $\xi_1 = O((\Delta x)^2)\ \tau_1 = O((\Delta x)^3)$ $\xi_{N-1} = O((\Delta x)^3)$ $\xi_i = O((\Delta x)^2)\sum\sum \tau_i = O((\Delta x)^2)$, $i = 2,\ldots,N-2$

and let Ω be any irregular 2-D shape inscribed within a rectangle with boundary $\partial\Omega$ at which Dirichlet conditions $\psi(x,y) = g(x,y)$ are specified. As $\psi = 0$ outside the physical domain, there may be jumps on $\partial\Omega$. We denote by $x_{i+1/2,j}$ the midpoint of the interval $[x_{i,j},x_{i+1,j}]$ (i.e., in the x-direction) and by $y_{i,j+1/2}$ the midpoint of the interval $[y_{i,j},y_{i,j+1}]$. The generalization of the numerical approach to the 2-D case is simple to implement since it involves a dimension by dimension application of the 1-D method. For all the interior points where there are no jumps in x, we use centered differences to approximate $\partial\psi/\partial x$ at the midpoint:

$$\psi'_{i+1/2,j} = \frac{\psi_{i+1,j} - \psi_{i,j}}{\Delta x}, \tag{35}$$

with a similar approach for $\partial\psi/\partial y$ at internal points. For any points where there is a jump in x to be incorporated in the discretization, we approximate $\partial\psi/\partial x$ at $x = x_{i+1/2,j}$ by fitting a linear polynomial through the values $\psi_{I,j}$ and $\psi_{i+1,j}$, and then evaluating its slope at $x = x_{i+1/2,j}$. Similarly, for any points where there is a jump in y, we need to approximate $\partial\psi/\partial y$ at $y = y_{i,j+1/2}$ by fitting a linear polynomial through the values $\psi_{i,J}$ and $\psi_{i,j+1}$, and evaluate its slope at $y = y_{i,j+1/2}$. This gives rise to a pentadiagonal stencil for the matrix representing the discrete operator, and with elementary modifications this can be made symmetric.

A similar approach is used for the quadratic boundary treatment, again dealing separately with x and y derivatives. However, for any points where there is a jump in x to be incorporated in the discretization, we approximate $\partial\psi/\partial x$ at $x = x_{i+1/2,j}$ by fitting a quadratic polynomial through the values $\psi_{I,j}$, $\psi_{i+1,j}$ and $\psi_{i+2,j}$, and then evaluating its slope at $x = x_{i+1/2,j}$. A similar treatment is used in the y-direction. Again, this gives rise to a pentadiagonal stencil for the matrix representing the discrete operator, but it is not possible to make this matrix symmetric. Either of the two boundary treatments naturally gives rise to three kinds of boundary points, those where a boundary treatment is needed for both x- and y-directions (corner points) and those where boundary treatment is only required for the x- or y-directions separately. All three kinds of points are shown for a circular domain embedded in a square box in Fig. 1. For example, the open square in Fig. 1(b) at $x \approx 0.55$ is a corner point, giving a boundary contribution to both the x-wise discretization at $y \approx -0.8$ and the y-wise discretization at $x \approx 0.55$. The adjacent open circle at $x \approx 0.61$ gives a

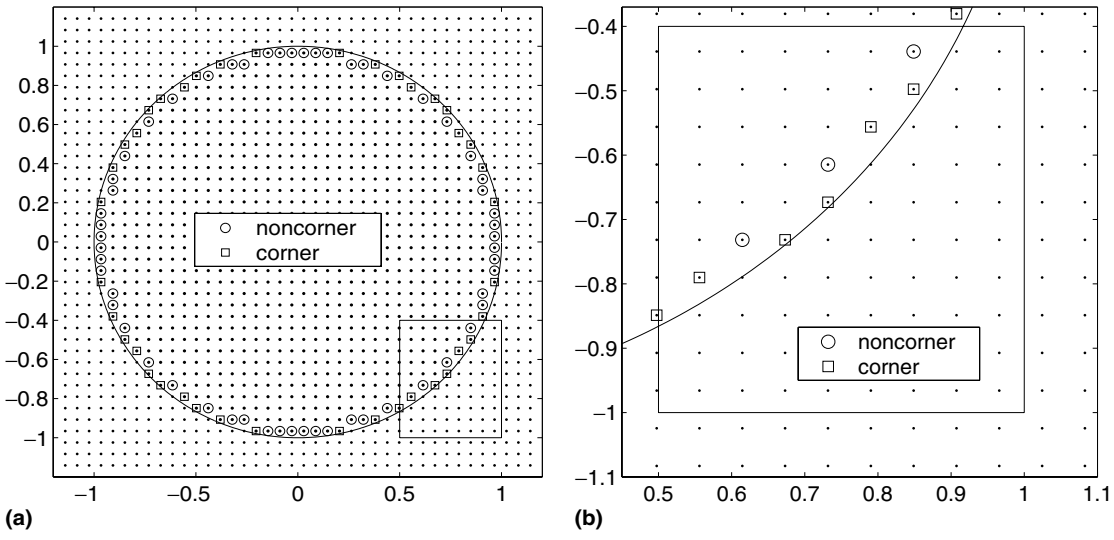

Fig. 1. Panel (a) shows the irregular region, in this case the unit circle, embedded in a square domain. The solid line is the real boundary, the open circles are the non-corner boundary points and the square boxes are the corner boundary points. Panel (b) is an enlargement of the box at the bottom right-hand corner of the circle in (a).

boundary contribution in the y-direction, but not in the x-direction, and the next open circle at $x \approx 0.73$ gives a boundary contribution only in the x-direction.

3. Numerical results

In this section, we give results for a selection of 1-D problems that show that the error expressions for the linear boundary treatment (Eq. (17) and the approximate forms (24) and (25)) and the quadratic boundary form (32) compare well with direct numerical calculations.

We then go on to show that the 1-D boundary expressions can be used to predict the error for 2-D problems, for boundary points that are not corner points. As the boundary error dominates for the linear treatment, this gives us a useful prediction for the maximum error in the 2-D case. By contrast, the boundary error when the quadratic boundary treatment is employed still shows agreement with the 1-D analytical expressions. In this case, the maximum error will occur internally, but is essentially equivalent to the error expected in the absence of boundary jumps for which we have no analytical expression.

3.1. 1-D cases

Fig. 2(a) shows the errors involved in using the 1-D linear boundary treatment of (3) for a symmetric problem, where $\psi = \cos x$. The domain of interest is embedded in the region $x \in [-1,1]$ with $\psi = 0$ for $|x| > 0.7$. Dirichlet conditions are applied at $x = \pm(0.7-0.3\pi/N)$ where there are N intervals of length $2/N$. Thus we are really solving a sequence of different problems, with the boundary condition dependent on N. This is done to ensure firstly that α_L and α_R are constant for all N so that convergence depends only on N: our earlier analysis shows strong dependence of the error on the magnitudes of α_L and α_R. (In fact here $\alpha_L = \alpha_R = 0.471$, which corresponds to almost the maximum error case $\alpha_L = \alpha_R = 1/2$.) The second aim is to guarantee that we have a test problem of relevance to the 2-D case, where the boundary conditions will not in general coincide with grid-points. (A possible alternative approach is to take $N-1$ intervals and fixed boundaries at $x = \pm 0.7$. This would satisfy the requirement that the boundaries not coincide with grid-points, but then α_L and α_R would vary with N.) We see almost exact agreement with the error expression (17). Similar good agreement between the quadratic boundary treatment of (4) and the full error expression (32) is shown in Fig. 2(b). Fig. 2(c) shows the rms and maximum absolute errors for the two cases: as would be expected from the form of the error in (a), the end-point error dominates the rms error as well as the maximum error, with the quadratic convergence clearly shown in all cases. (Results are shown for $N = 40$, 80, 160 and 320 and plotted as a function of $\log_{10}(N)$.) Finally in Fig. 2(d), we see that the sum of approximate forms for the end-point errors (22) and (23) gives a very useful estimate for the dominant error, and that once the internal error expression (25) is added the approximate form is essentially identical to the full error expression. Fig. 3 is similar to Fig. 2, but for the case with analytical solution $\psi = e^x$, which is no longer symmetric about $x = 0$. We find qualitatively similar results, with the linear boundary treatment, dominated by the error at the boundaries, giving rise to somewhat larger errors overall than in the case of Fig. 2. The final 1-D example of Fig. 4 has $\psi = \sin 3x$. Here there is some cancellation between the left- and right-hand errors, so that the ratio of linear to quadratic error is significantly less than in the two previous cases, with ratio $r \approx 3$ for both rms and maximum absolute errors.

3.2. 2-D cases

We now discuss the application of the method to 2-D problems. In Fig. 5 we consider the problem of inverting the Poisson equation $\nabla^2\psi = -2\cos(x+y)$ on the unit circle, with Dirichlet boundary conditions chosen to ensure continuity from within, i.e., $\psi = \cos(x+y)$. Outside the unit circle we set $\psi = 0$. The

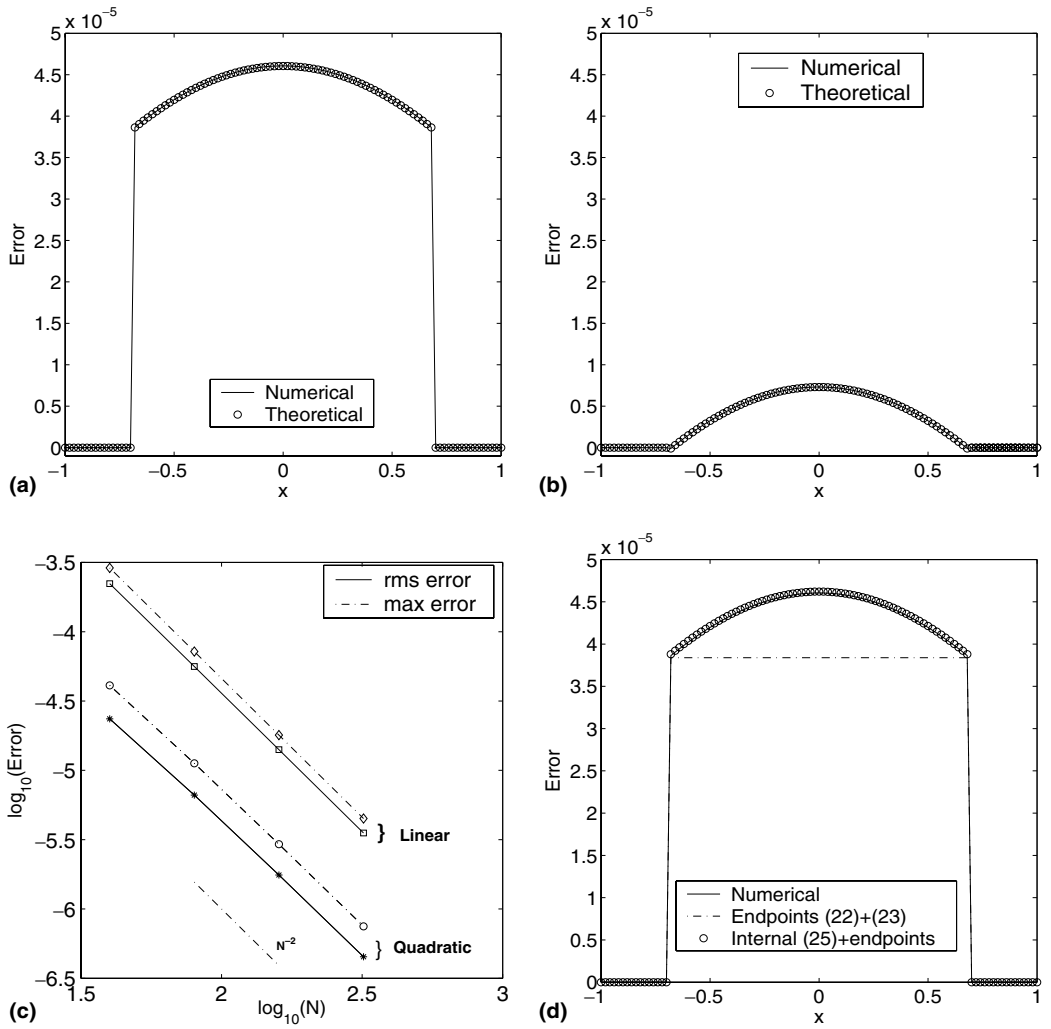

Fig. 2. The error in solving the 1-D problem $d^2\psi/dx^2 = -\cos x$, with boundary conditions $\psi = \cos[\pm(0.7 - 0.3\pi/N)]$ for $x = \pm(0.7 - 0.3\pi/N)$, respectively, and with $\psi = 0$ elsewhere. In (a) the numerical results for $N = 100$ are shown with a solid curve for the linear boundary treatment of (3), compared with the full analytic error expression (17), shown as open circles. In (b) the quadratic boundary treatment (4) with $N = 100$ (solid curve) is compared with the quadratic error expression (32) (open circles). Panel (c) compares the rms error (solid lines) and maximum absolute error (dash-dotted lines) for the two cases. Finally, in panel (d) we compare the numerical error for the linear case of (a) with the sum of the end-point errors calculated using (22) and (23) (dash-dotted curve). Also shown (as open circles) is the sum of the end-point errors (22) and (23) and the integral form for the internal error (25).

domain is embedded in a square of side length 2.4. The first panel (a) shows a contour plot of the solution ψ, with the discontinuity at the boundary clearly apparent and using $N = 40$, where N is the number of grid-intervals in each coordinate direction. Panel (b) shows the rms and maximum errors as a function of $\log_{10} N$ for $N = 20, 40, 80$ and 160. All errors scale with N^{-2}, as for the 1-D case, with the quadratic boundary treatment reducing the rms error by a factor of about 3.4. In panels (c) and (d), we show the boundary error for the linear and quadratic boundary treatments, respectively (again with $N = 40$, so that there are 96 boundary points just inside the circle, as marked with open circles and squares in Fig. 1). As we only have available the 1-D error expressions, we can only make comparisons at points on the boundary which essentially

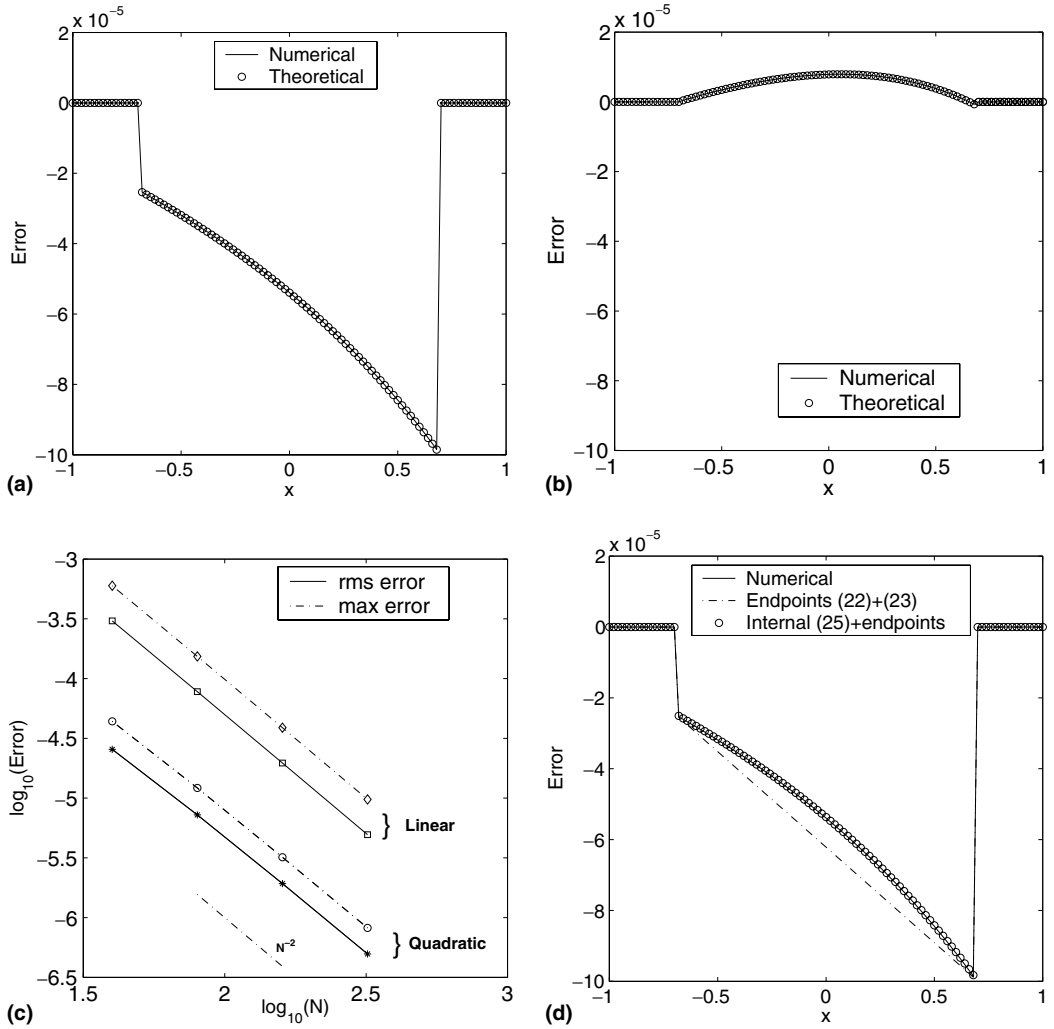

Fig. 3. As for Fig. 2 but with the 1-D problem $d^2\psi/dx^2 = e^x$, with boundary conditions $\psi = e^{\pm(0.7 - 0.3\pi/N)}$ for $x = \pm(0.7 - 0.3\pi/N)$, respectively.

use a 1-D treatment, i.e., boundary points which are not corner points (see Fig. 1). Although the agreement is not exact, we see quite reasonable estimation of the error at the boundaries, with the boundary error for the quadratic treatment 35 times smaller than for the linear treatment. The numerical errors for the whole domain are plotted in Fig. 6 (panels (a) and (c) for the linear treatment, and (b) and (d) for the quadratic treatment). As in the 1-D problems, when the linear treatment is used the maximum error occurs at the boundaries. Therefore our 1-D estimation of boundary error gives a reasonable bound for the error over the whole domain in this case. By contrast, the quadratic error is dominated by the internal error (as would be found in the absence of boundary jumps) and the boundary contribution is negligible. This confirms the superiority of the quadratic boundary treatment. A more complicated domain and form for ψ are considered in Figs. 7 and 8, with similar features apparent. In fact in this case, the linear boundary error is even larger by comparison with the quadratic boundary error (note that in Fig. 8(d) the errors have been multiplied by a factor of 10.)

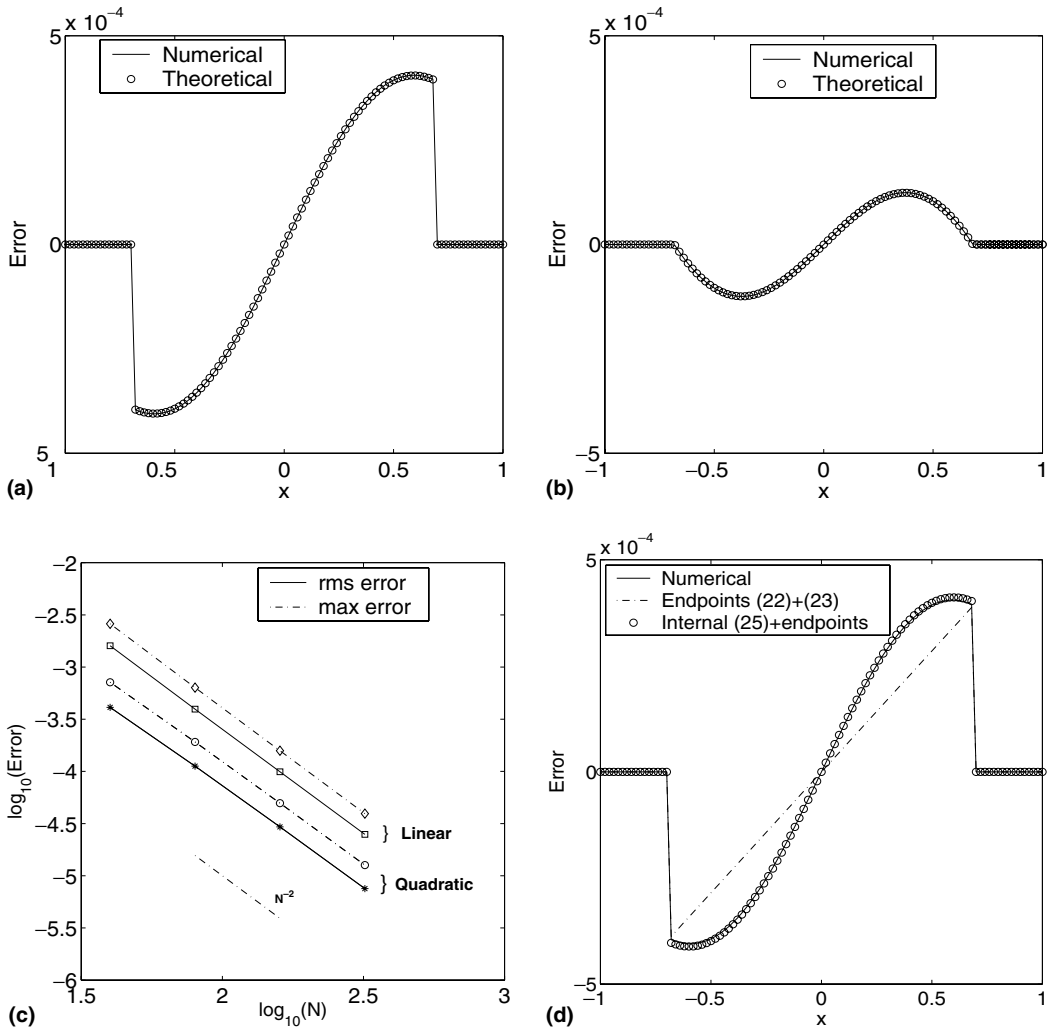

Fig. 4. As for Fig. 2 but with the 1-D problem $d^2\psi/dx^2 = -9\sin 3x$, with boundary conditions $\psi = \sin[\pm 3(0.7 - 0.3\pi/N)]$ for $x = \pm(0.7 - 0.3\pi/N)$, respectively.

The method of [9] also uses a quadratic treatment at the boundaries, but does not treat the x- and y-directions separately, but rather uses an area fit for the boundary partial cells. They show that this leads to $O((\Delta x)^3)$ error from the boundaries, with $O((\Delta x)^2)$ internal error dominating for large N, as is the case with the present method. In Fig. 9, we make comparison with a case treated in [9]. They consider a six-leaf domain, similar to the four-leaf domain of Fig. 7, described by $r \leqslant 0.30 + 0.15\cos 6\theta$ with $\nabla^2\psi = 7r^2\cos 3\theta$ on the interior of the domain and non-zero Dirichlet boundary conditions chosen to ensure continuity with the interior solution, i.e., $\psi(r,\theta)=r^4\cos 3\theta$. Using data from their Table 2 and their Fig. 9, for the one-norm defined as the mean absolute error, shows that the coefficient of error from the boundary terms is smaller with the Shortley–Weller approach, so that $O((\Delta x)^2$ behaviour holds even for moderate N. For large N, both methods appear to have the same order of error, but the method used here appears slightly more efficient. We expect this behaviour to hold in general, as it is consistent with all our other 2-D cases.

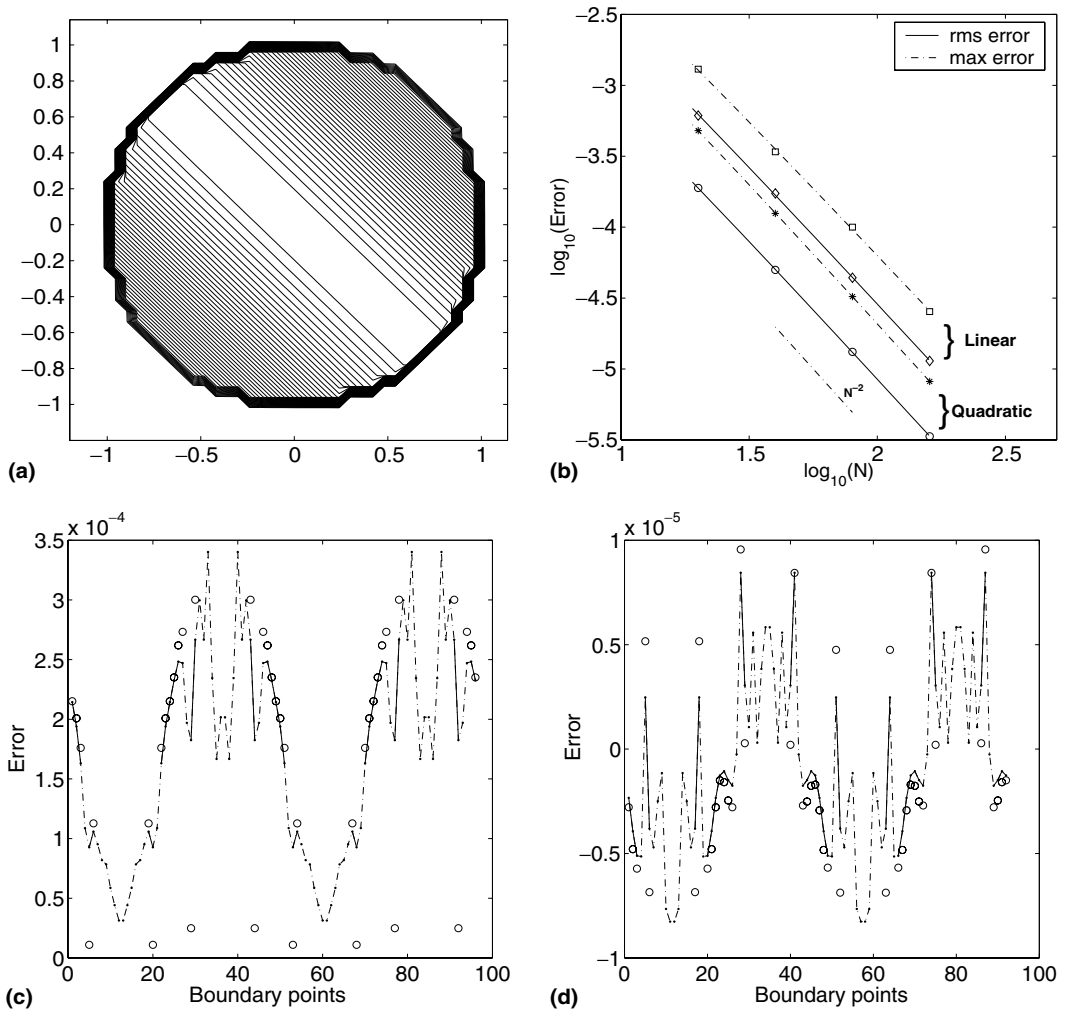

Fig. 5. The 2-D problem $\nabla^2\psi = -2\cos(x + y)$, with $N\Delta x = 2.4$ and $N = 40$ (panels (a), (c) and (d)), subject to $\psi = \cos(x + y)$ on the boundary of the unit circle. In (a) the numerical results are shown contoured for the quadratic boundary treatment. (The linear results look very similar.) Panel (b) compares the rms error (solid lines) and maximum absolute error (dash-dotted lines) for the quadratic (Eq. (4)) and linear (Eq. (3)) cases. In (c) the actual numerical errors for the linear boundary treatment are shown at all points on the boundary. These are taken as running clockwise from $(x,y) = (-1,0)$ (labelled 1 to 96) and the errors are shown as a dash-dotted curve. Where there are sequences of adjacent non-corner points, the dash-dotted line for the numerical error is replaced with a solid line. For non-corner points, the 1-D approximate linear errors of Eqs. (22) and (23) (in the x-direction) and the equivalent in the y-direction are shown as open circles. Finally, in panel (d) we show the numerical error for quadratic boundary treatment, using the same labelling and format as in (c). The open circles now represent (for non-corner points) the 1-D full quadratic error of (32) for the x-direction and the equivalent in the y-direction.

Note that in all the above calculations, the corresponding matrix problems have been solved by direct inversion. In application of these ideas in ongoing work on vortex dynamics on an irregular 2-D domain using the CASL technique (i.e., extensions of the circular cylinder case of [12]) we store the LU decomposition of the matrix at the beginning of the calculation. Extension to the full quasi-geostrophic (layerwise 2-D) case will require storage of the corresponding decomposition for each vertical mode. For very large

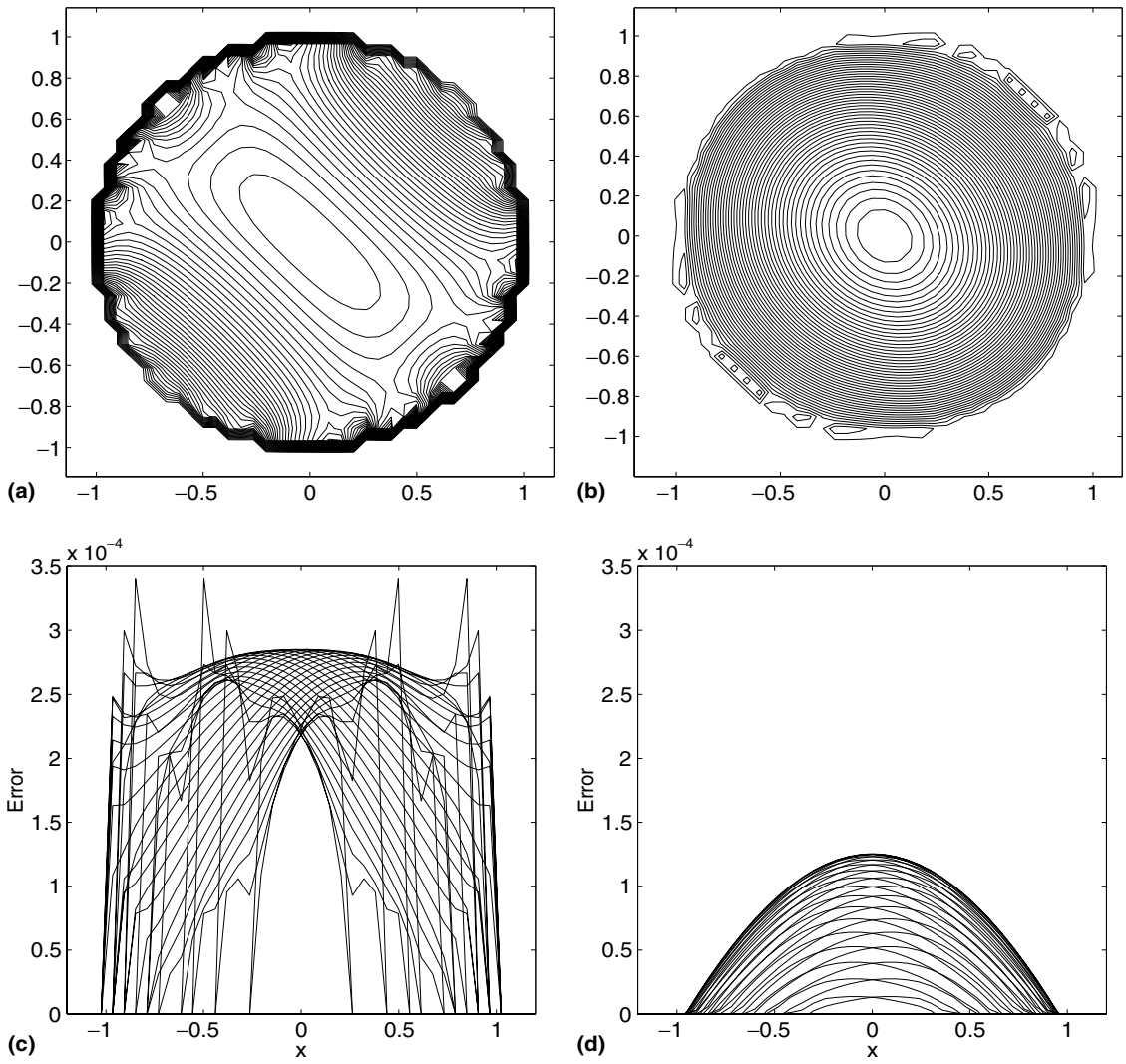

Fig. 6. The 2-D problem $\nabla^2\psi = -2\cos(x + y)$, with $N\Delta x = 2.4$ and $N = 40$, subject to $\psi = \cos(x + y)$ on the boundary of the unit-circle (as in Fig. 5): (a) contour plot of the error with the linear boundary treatment, (b) contour plot of the error with the quadratic boundary treatment, (c) plots of error at constant values of y for the linear boundary treatment and (d) plots of error at constant values of y for the quadratic boundary treatment.

problems, storage limitations may require a shift to an iterative solution of each matrix problem at each time-step, but this will have a significant time penalty.

4. Conclusion

In this paper, we have discussed the embedding method for the solution of the Poisson equation, using both a linear and a quadratic boundary treatment following [3,21], respectively. For the 1-D case, we have

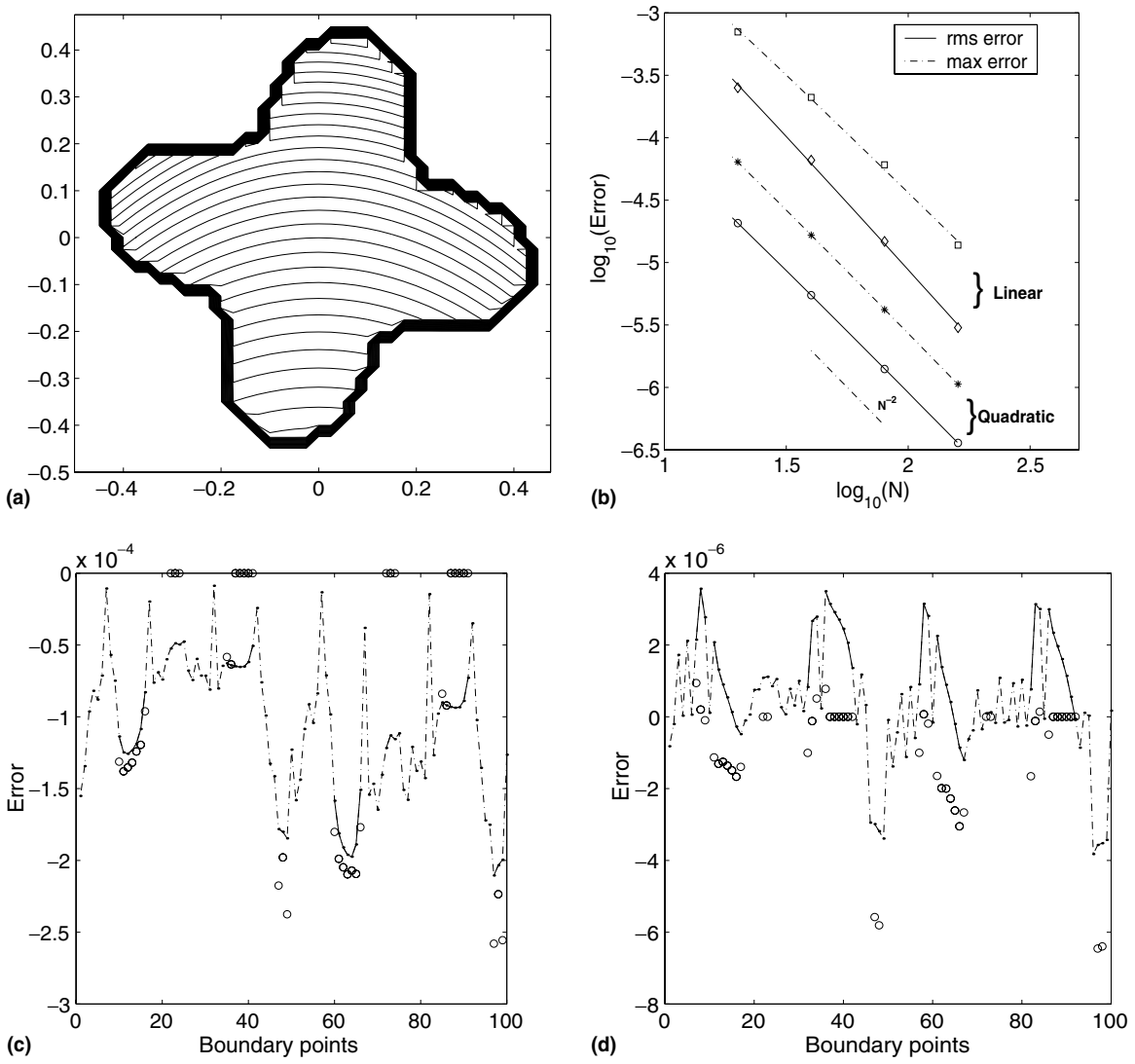

Fig. 7. As in Fig. 5 but for the 2-D problem $\nabla^2\psi = (3 + 4x^2)\exp(x^2 + y)$, $N\Delta x = 1$ and $N = 40$, subject to $\psi = \exp(x^2 + y)$ on the boundary of the 'rotated four-leaf shape' described by $r = 0.35 + 0.1\cos[4(\theta + \pi/17)]$.

determined explicit forms for the error involved, following the ideas of [9]. Furthermore, for the linear boundary treatment we have derived simple approximate expressions that show how the dominant error is related to the end-point truncation error. These expressions show that for the linear case the boundary error is $O((\Delta x)^2)$ in agreement with [3,7] and so is of the same order as the internal error that is present even in the absence of boundary jumps. However, the coefficient of error is very large for the boundary contribution and so usually dominates. Similarly, in agreement with [21,9] we find that the end-point error is $O((\Delta x)^3)$ for the quadratic boundary treatment and so has negligible influence, with the total error being dominated by the internal error.

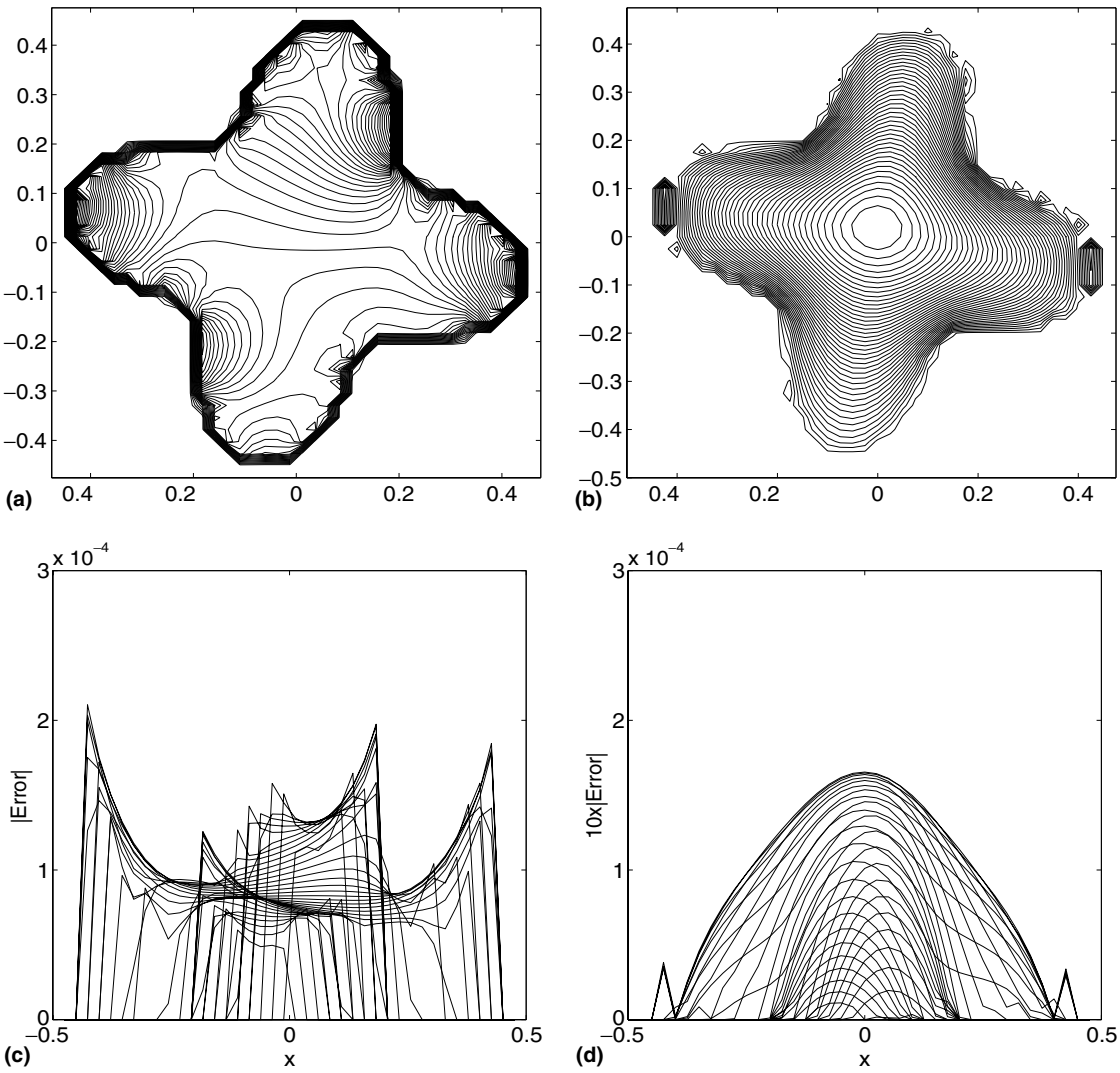

Fig. 8. The 2-D problem $\nabla^2\psi = (3 + 4x^2)\exp(x^2 + y)$, with $N\Delta x = 1$ and $N = 40$, subject to $\psi = \exp(x^2 + y)$ on the boundary of the 'rotated four-leaf shape' described by $r = 0.35 + 0.1\cos[4(\theta + \pi/17)]$ (as in Fig. 7): (a) contour plot of the error with the linear boundary treatment, (b) contour plot of the error with the quadratic boundary treatment, (c) plots of absolute error at constant values of y for the linear boundary treatment and (d) plots of absolute error (multiplied by ten for ease of visualization) at constant values of y for the quadratic boundary treatment.

Although no expressions of error have been derived for the 2-D case, we have found that for boundary points that are not corner points, the 1-D error expressions, in either the x- or y-directions as appropriate, give good estimates for the 2-D boundary error in both linear and quadratic cases. As the boundary error dominates for the linear boundary treatment, this means that we have useful bounds for the total error in this case (and indeed we expect this to generalise to 3-D, although no calculations have been performed). By contrast for the quadratic case we find that the boundary error is negligible.

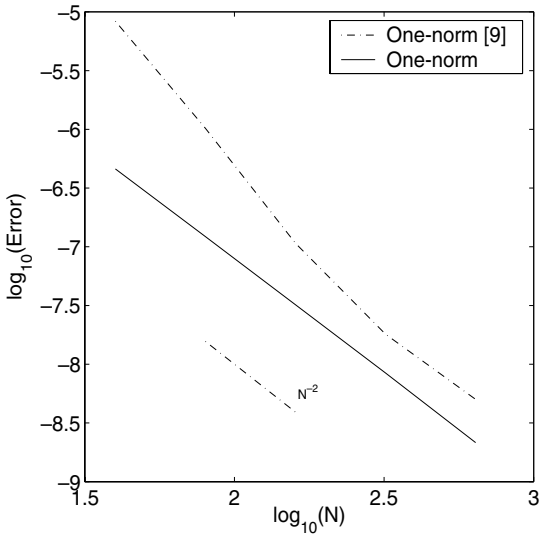

Fig. 9. Comparison of the Shortley–Weller method used here (solid curve) with the results from table 2 and Fig. 9 of [9] (shown dot-dashed) for the one-norm of the error (mean absolute error). The domain is a six-leaf figure – see the main text for details.

Thus in summary we have shown that the 1-D error treatment gives significant information about the 2-D Poisson solution method, particularly for the linear boundary treatment. In particular, we have demonstrated that if a non-symmetric matrix formulation can be accepted, then the quadratic boundary treatment should always be employed. We believe this conclusion will also apply to the 3-D problem, as the boundary treatment is carried out dimension by dimension.

We note that because the expressions for end-point error (22) and (23) take such a simple form in 1-D it is possible to correct for this error explicitly in the matrix formulation. If this is done, the quadratic boundary treatment is exactly recovered, with the attendant loss of symmetry.

References

[1] K. Bohmer, Asymptotic error expansions and discrete Newton methods for elliptic boundary-value-problems, Lect. Notes Math. 968 (1982) 292–300.
[2] J.H. Bramble, B.E. Hubbard, On the formulation of finite difference analogues of the Dirichlet problem for Poisson's equation, Numer. Math. 4 (1962) 313–327.
[3] L. Collatz, Bemerkungen zur fehlerabschätzung für das differnzenverfahren bei partiellen differentialgleichungen, Z. Angew. Math. Mech. 13 (1933) 56–57.
[4] L. Collatz, The Numerical Treatment of Differential Equations, Springer, Berlin, 1966.
[5] Q. Fang, T. Yamamoto, Superconvergence of finite difference approximations for convection–diffusion problems, Numer. Linear Algebra Appl. 8 (2001) 99–110.
[6] S. Gerschgorin, Fehlerabschätzung für das differnzenverfahren zur lösung partieller differentialgleichungen, Z. Angew. Math. Mech. 10 (1930) 373–382.
[7] F. Gibou, R.P. Fedkiw, L.-T. Cheng, M. Kang, A second-order accurate symmetric discretization of the Poisson equation on irregular domains, J. Comp. Phys. 176 (2002) 205–227.
[8] D.W. Hewett, The embedded curved boundary method for orthogonal simulation meshes, J. Comp. Phys. 138 (1997) 585–616.
[9] H. Johansen, P. Colella, A Cartesian grid embedded boundary method for Poisson's equation on irregular domains, J. Comp. Phys. 147 (1998) 60–85.
[10] R.J. LeVeque, Z. Li, The immersed interface method for elliptic equations with discontinuous coefficients and singular sources, SIAM J. Numer. Anal. 31 (1994) 1019–1044.

[11] X. Liu, R. Fedkiw, M. Kang, A boundary condition capturing method for Poisson's equation on irregular domains, J. Comp. Phys. 160 (2000) 151–178.
[12] C. Macaskill, W.E.P. Padden, D.G. Dritschel, The CASL algorithm for quasi-geostrophic flow in a cylinder, J. Comp. Phys. 188/1 (2003) 232–251.
[13] N. Matsunaga, T. Yamamoto, Superconvergence of the Shortley–Weller approximation for Dirichlet problems, J. Comput. Appl. Math. 116 (2000) 263–273.
[14] A. Mayo, The fast solution of Poisson's and the biharmonic equations in irregular domains, SIAM J. Numer. Anal. 21 (2) (1984) 285–299.
[15] A. Mayo, The rapid evaluation of volume integrals of potential theory on general regions, J. Comp. Phys. 100 (1992) 236–245.
[16] A. Mayo, A. Greenbaum, Fast parallel iterative solution of Poisson's and the biharmonic equations on irregular regions, SIAM J. Sci. Stat. Comput. 13 (1) (1992) 101–117.
[17] A. McKenney, L. Greengard, A. Mayo, A fast Poisson solver for complex geometries, J. Comp. Phys. 118 (1995) 348–355.
[18] V. Pereyra, W. Proskurowski, O. Widlund, High order fast Laplace solvers for the Dirichlet problem on general regions, Math. Comp. 31 (1977) 1–16.
[19] C.S. Peskin, B.F. Printz, Improved volume conservation in the computation of flows with immersed elastic boundaries, J. Comp. Phys. 105 (1993) 33–46.
[20] C.S. Peskin, Numerical analysis of blood flow in the heart, J. Comp. Phys. 25 (1977) 220–252.
[21] G.H. Shortley, R. Weller, The numerical solution of Laplace's equation, J. Appl. Phys. 9 (1938) 334–348.
[22] H.S. Udaykumar, R. Mittal, W. Shyy, Computation of solid–liquid phase fronts in the sharp interface limit on fixed grids, J. Comp. Phys. 153 (1999) 535–574.
[23] W. Wasow, Discrete approximations to elliptic differential equations, Z. Angew. Math. Phys. 6 (1955) 81–97.