数理解析研究所講究録
1040 巻 1998 年 135-142

On the Accuracy of Finite Difference Solution for Dirichlet Problems

Tetsuro Yamamoto(山本 哲朗)
Department of Mathematical Sciences
Ehime University
Matsuyama 790-8577, Japan

1. Introduction

Let Ω be a bounded domain of $\mathbf{R}^2$ and consider the Dirichlet problem

$$\begin{cases} -\Delta u+c(x,y)u = f(x,y) & \text{in} \quad \Omega & (1.1) \\ u = g(x,y) & \text{on} \quad \Gamma=\partial\Omega & (1.2) \end{cases}$$

where c,f and g are given functions satisfying $c\geq 0$,

$$c,f\in C^{0,\alpha}(\Omega)=C^{\alpha}(\Omega) \quad \text{and} \quad g\in C(\Gamma)$$

with the Hölder exponent $\alpha\in(0,1)$. Then it is known [2,4] that there exists a unique solution $u\in C(\bar{\Omega})\cap C^{2,\alpha}(\Omega)$ of (1.1) and (1.2). Furthermore, if l is a nonnegative integer,

$$c,f\in C^{l,\alpha}(\bar{\Omega}), \quad g\in C^{l+2,\alpha}(\bar{\Omega})$$

and Ω is a $C^{l+2,\alpha}$ domain, then it is also known that

$$u\in C^{l+2,\alpha}(\bar{\Omega}). \tag{1.3}$$

Finite difference methods for solving the problem (1.1)-(1.2) have extensively been studied in much literature (e.g., [3],[5-6],[8-9],[12]) usually for the case $u\in C^{4}(\bar{\Omega})$. We can find there many estimates on the accuracy of finite difference formulas. The accuracy of the formula, however, does not necessarily imply that of the approximate solution. Furthermore, it appears to the author that there is no explicit mention about superconvergence of discretized solution in any literature.

In this paper, we shall first give a convergence theorem for the Shortley-Weller discretization, which also asserts a superconvergence of the discretized solution near the boundary Γ. Furthermore, our argument can be applied to the equations in polar coordinate systems to obtain the similar result. Finally we point out that the argument can also be applied to a Dirichlet problem of a semilinear equation of the form

$$-\Delta u+f(x,y,u)=0,$$

where $f \in C^2(\bar{\Omega} \times \mathbf{R})$ and $\dfrac{\partial f}{\partial u} \geq 0$ in $\bar{\Omega} \times \mathbf{R}$.

Throughout this paper, we put $C^{l,0}(\bar{\Omega}) = C^l(\bar{\Omega})(C^{l,0}(\Omega) = C^l(\Omega))$ and use the notation $C^{l,\alpha}(\bar{\Omega})(C^{l,\alpha}(\Omega))$ as the set of functions whose l-th order partial derivatives are Hölder (locally Hölder) continuous in $\bar{\Omega}(\Omega)$. Recall that u is called Hölder continuous with exponent $\alpha(0 < \alpha < 1)$ in a domain D if

$$\sup_{\substack{R,Q\in D\\ P\neq Q}} \frac{|u(P)-u(Q)|}{||P-Q||^\alpha} < \infty,$$

where $||\cdot||$ stands for the Euclidean norm, and locally Hölder continuous in D if u is Hölder continuous on any compact subset of D. This definition is extended to the case $\alpha = 1$, where "Hölder (or locally Hölder)" is replaced by "Lipschitz (or locally Lipschitz)."

2. Accuracy of the Shortley-Weller Approximation

Let $h = \Delta x$ and $k = \Delta y$ be the mesh sizes in x, y directions and put

$$x_i = x_{i-1} + h, \quad y_j = y_{j-1} + k, \quad i = 1, 2, \cdots, \mathrm{I} \quad, j = 1, 2, \cdots, \mathrm{J}.$$

The grid point (x_i, y_j) in Ω is often written as P_{ij}. We shall say that the point P_{ij} is near Γ if the distance $d(P_{ij}, \Gamma)$ between P_{ij} and Γ is at most $O(h+k)$. P_{ij} is called a quasi-boundary point if at least one of the four points $(x_i \pm h, y_j), (x_i, y_j \pm k)$ does not belong to $\bar{\Omega} = \Omega \cup \Gamma$. Otherwise, P_{ij} is called a normal (grid) point. We denote by $\mathcal{P}_0$ and $\mathcal{P}_\Gamma$ the set of normal points and the set of quasi-boundary points, respectively and put $\Omega_{hk} = \mathcal{P}_0 \cup \mathcal{P}_\Gamma$.

Let the four neighbor points of $P \in \Omega_{hk}$ be denoted by P_E, P_W, P_S and P_N and their distances to P be denoted by h_E, h_W, k_S and k_N, respectively (cf. Figure 1).

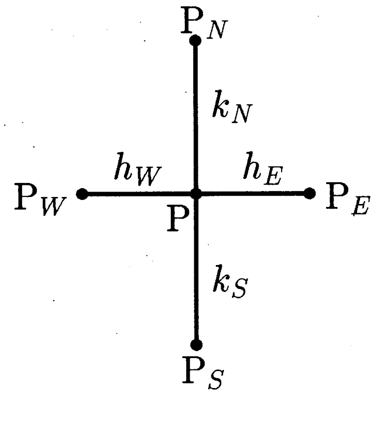

Fig. 1

Furthermore, we denote by $U(P)$ the finite difference solution at P. Then the Shortley-Weller (S-W) approximation (cf. [3], [5]) for $-\Delta u$ is a five point formula defined by

$$\begin{aligned} -\Delta_{hk}U(P) &= \left(\frac{2}{h_E h_W}+\frac{2}{k_S k_N}\right)U(P) \\ &\quad -\frac{2}{h_E(h_E+h_W)}U(P_E)-\frac{2}{h_W(h_E+h_W)}U(P_W) \\ &\quad -\frac{2}{k_S(k_S+k_N)}U(P_S)-\frac{2}{k_N(k_S+k_N)}U(P_N), \end{aligned} \tag{2.1}$$

which reduces to the usual five point formula if $h_E = h_W = h$ and $k_S = k_N = k$.

If $u \in C^4(\bar{\Omega})$, then the truncation error of u at P is given by

$$\begin{aligned} \tau(P) &\equiv -(\Delta_{hk}u(P)-\Delta u(P)) \\ &= \frac{h_E-h_W}{3}u_{xxx}(P)+\frac{k_N-k_S}{3}u_{yyy}(P) \\ &\quad +\frac{h_E^2-h_Eh_W+h_W^2}{12}u_{xxxx}(Q_H)+\frac{k_S^2-k_Ek_W+k_N^2}{12}u_{yyyy}(Q_V) \\ &= \begin{cases} O(h^2+k^2) & (\text{if} \quad P\in\mathcal{P}_0) \qquad (2.2)\\ O(h+k) & (\text{if} \quad P\in\mathcal{P}_\Gamma), \qquad (2.3) \end{cases} \end{aligned}$$

where Q_H and Q_V are points on the lines $\overline{P_W P_E}$ and $\overline{P_N P_S}$, respectively.(Note that (2.2) and (2.3) also hold for the case $u \in C^{3,1}(\bar{\Omega})$.)

Similarly, if $u \in C^{l+2,\alpha}(\bar{\Omega})$, then we have

$$\tau(P) = \begin{cases} O(h^{l+\alpha}+k^{l+\alpha}) & (\text{if} \ \ p\in\mathcal{P}_0 \ \ \text{and} \ \ l=0 \ \ \text{or} \ \ 1) \qquad (2.4)\\ O(h^\alpha+k^\alpha) & (\text{if} \ \ p\in\mathcal{P}_\Gamma \ \ \text{and} \ \ l=0) \qquad (2.5)\\ O(h+k) & (\text{if} \ \ p\in\mathcal{P}_\Gamma \ \ \text{and} \ \ l=1). \qquad (2.6) \end{cases}$$

Let N be the number of the grid points P_{ij} in Ω and arrange them as $P_1, \cdots, P_N$ in appropriate order. We then put

$$\tau = (\tau(P_1),\cdots,\tau(P_N))^t = (\tau_1,\cdots,\tau_N)^t,$$

$$\mathbf{U} = (U(P_1),\cdots,U(P_N))^t = (U_1,\cdots,U_N)^t,$$

$$\mathbf{u} = (u(P_1),\cdots,u(P_N))^t = (u_1,\cdots,u_N)^t$$

and

$$C = \text{diag}(c_1,\cdots,c_N),$$

where $c_i = c(P_i)$. Then the vectors $\mathbf{U}$ and $\mathbf{u}$ satisfy the following systems of linear equations

$$(A+C)\mathbf{U} = \mathbf{b}$$

and

$$(A+C)\mathbf{u} = \mathbf{b}+\tau$$

where $A = (a_{ij})$ is an $N \times N$ irreducibly diagonally dominant L-matrix and b is an N-dimensional vector which comes from the boundary condition (1.2). Recall that a matrix A is called an L-matrix if $a_{ii} > 0$ and $a_{ij} \leq 0 (i \neq j)$ (cf. [13]) and that an irreducibly diagonally dominant L-matrix is an M-matrix.

We then have

$$(A + C)(\mathbf{u} - U) = \tau. \tag{2.7}$$

This implies

$$u - \mathbf{U} = (A + C)^{-1}\tau$$

and

$$|u - \mathbf{U}| \leq (A + C)^{-1}|\tau| \leq A^{-1}|\tau| \leq ||\tau||_\infty A^{-1}\mathbf{e} \tag{2.8}$$

where we put

$$\begin{aligned} |u - \mathbf{U}| &= (|u_1 - U_1|, \cdots, |u_N - U_N|)^t, \\ |\tau| &= (|\tau_1|, \cdots, |\tau_N|)^t, \\ \mathbf{e} &= (1, \cdots, 1)^t, \end{aligned}$$

and we have used the fact that $A + C$ is an M-matrix and $0 \leq (A + C)^{-1} \leq A^{-1}$ since C is a nonnegative matrix (cf. [12]). Hence, estimating $A^{-1}\mathbf{e}$ in the right-hand side of (2.8) yields error bounds for the finite difference solution. This technique can be found in Varga [12] and Ortega [6], and extended arguments are found in Hackbush [5], where it is assumed, however, that $u \in C^4(\bar{\Omega})$ or $C^{3,1}(\bar{\Omega})$ and $h = k$. More sophisticated analysis along this line leads to the following result.

Theorem 1 (Superconvergence of the S-W Approximation).
(i) If $u \in C^{3,1}(\bar{\Omega})$, then

$$|u(P) - U(P)| \leq \begin{cases} O(h^2 + k^2) & (P \in \mathcal{P}_0) \\ O(h^3 + k^3) & (P \text{ is near } \Gamma) \end{cases}$$

(ii) If $u \in C^{l+2,\alpha}(\bar{\Omega})$, $l = 0$ or 1 and $0 < \alpha < 1$, then

$$|u(P) - U(P)| \leq \begin{cases} O(h^{l+\alpha} + k^{l+\alpha}) & (P \in \mathcal{P}_0) \\ O(h^{l+1+\alpha} + k^{l+1+\alpha}) & (P \text{ is near } \Gamma) \end{cases}$$

Remark 2.1. For the S-W approximation, the truncation error at every quasi- boundary point is at most $O(h + k)$ if $u \in C^{3,1}(\bar{\Omega})$. Nevertheless, Theorem 1 shows the third-order accuacy of the finite difference solution at the points near Γ and the second order accuracy at other points.

3. Accuracy of the Swartztrauber-Sweet Approximation in Polar Coordinate Systems

If Ω is the open disk $\{(x,y)|x^2+y^2<R^2\}$ where R is a positive constant, then the problem (1.1)-(1.2) is usually solved by tranforming into the polar coordinate systems :

$$\begin{cases} -\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 u}{\partial\theta^2}\right]+c(r,\theta)u=f(r,\theta), & 0<r<R, 0\le\theta<2\pi \\ u=g(\theta), & r=R, 0\le\theta\le 2\pi. \end{cases} \tag{3.1}$$

According to Swartztrauber-Sweet [11], we discretize this as follows :

$$h=\Delta r=\frac{R}{m+1}\quad,\quad r_i=ih,\ i=0,\frac{1}{2},1,\cdots,m+\frac{1}{2},m+1 \tag{3.2}$$

$$k=\Delta\theta=\frac{2\pi}{n}\quad,\quad \theta_j=jk,\ j=0,1,2,\cdots,n-1,n \tag{3.3}$$

$$-\left[\frac{1}{r_ih^2}\left\{r_{i+\frac{1}{2}}(U_{i+1j}-U_{ij})-r_{i-\frac{1}{2}}(U_{ij}-U_{i-1j})\right\}\right.$$
$$\left.+\ \frac{1}{r_i^2k^2}(U_{ij+1}-2U_{ij}+U_{ij-1})\right]+c_{ij}U_{ij}=f_{ij}, \tag{3.4}$$
$$i=1,2,\cdots,m,\quad j=0,1,2,\cdots,n-1$$

$$U_{in}=U_{i0}\ \ (\forall i), U_{0j}=U_{00}\ \ (\forall j) \tag{3.5}$$

where U_{ij} stand for approximate solutions at $P_{ij}=(r_i,\theta_j)$. At the origin, we employ the formula

$$\left(1+\frac{c_{00}}{4}\right)U_{00}-\frac{1}{n}\sum_{j=0}^{n-1}U_{ij}=\frac{h^4}{4}f_{00}, \tag{3.6}$$

whose truncation error is $\tau_{00}=O(h^4)+o(k^4)$. For the case $c=0$, they proposed the scheme (3.2)-(3.6) without any convergence proof. Furthermore, in 1986, Strikwerda-Nagel [10] showed the second-order accuracy of the scheme by numerical experiments, but with no proof. It appears that any convergence proof for the above scheme has not been given since then.

For the Swartztrauber-Sweet (S-S) approximation (3.2)-(3.6), we have the following superconvergence result :

Theorem 2 (Superconvergence of the S-S Approximation).
(i) If $u\in C^{2,\alpha}(\bar{\Omega})$ with $\alpha\in(0,1)$, then

$$|u(P)-U(P)|\le\begin{cases} O(h^\alpha+k^\alpha) & (P\in\mathcal{P}_0) \\ O((h^\alpha+k^\alpha)h) & (\mathrm{dis}(P,r=R)=O(h)) \end{cases}$$

(ii) If $u \in C^{3,1}(\bar{\Omega})$, then

$$|u(P) - U(P)| \leq \begin{cases} O(h^2 + k^2) & (P \in \mathcal{P}_0) \\ O(h^3 + k^2 h) & (\text{dis}(P, r = R) = O(h)) \end{cases}$$

Remark 3.1. In [6], adding the condition

$$\lim_{r \to 0} r \frac{\partial u}{\partial r} = 0,$$

Samarsky-Andreev have considered another scheme for solving (3.1) with $c = 0$:

$$h > 0, \quad r_i = (i + \frac{1}{2})h, \qquad i = 0, 1, 2, \cdots, m + 1, \tag{3.7}$$

$$k = \frac{2\pi}{n}, \quad \theta_j = jk, \qquad j = 0, 1, 2, \cdots, n - 1, \ n \ , \rho(r) = r - \frac{h}{2} \tag{3.8}$$

$$-\left[\frac{1}{r_i}\left(\rho_{i+1}\frac{U_{i+1j} - U_{ij}}{h} - \rho_i \frac{U_{ij} - U_{i-1j}}{h}\right) + \frac{1}{r_i}\frac{1}{k^2}(U_{ij+1} - 2U_{ij} + U_{ij-1})\right] = f_{ij} \quad (i \geq 1) \tag{3.9}$$

$$-\left[\frac{1}{r_0 h}(U_{1j} - U_{0j}) + \frac{1}{r_0^2}\frac{1}{k^2}(U_{0j+1} - 2U_{0j} + U_{0j-1})\right] = f_{0j} \quad (i = 0), \tag{3.10}$$

where $\rho_i = \rho(r_i)$. With the use of the maximum principle, they proved

$$|u_{ij} - U_{ij}| \leq O(h^2 + k^2), \quad \forall i, j.$$

We remark here that Theorem 2 holds true for the scheme (3.7)-(3.10), too.

Remark 3.2. In [1], Chen considered asymptotic behavior of finite difference approximation for a radially symmetric solution $u = u(r)$ of a quasilinear parabolic equation

$$\frac{\partial u}{\partial t} = \Delta u + u^{1+\lambda}, \quad (t, x) \in (0, T) \times \Omega,$$

where Ω is an N-dimensional ball. He proved the $O(h^2)$-convergence of his scheme which discretizes $\frac{\partial^2 u}{\partial r^2}$ and $\frac{\partial u}{\partial r}$ in $\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{N-1}{r}\frac{\partial u}{\partial r}$ with the use of the centered difference. It is easy to see that a superconvergence result similar to Theorem 2 holds in this case, too.

4. Final Comments

(i) If Ω is a rectangle, then the smoothness of the solution will generally decrease at corners. However, some conditions are known for guaranteeing $u \in C^{3,\alpha}(\bar{\Omega}), C^{5,\alpha}(\bar{\Omega})$, etc. For such cases, the similar superconvergence property as in Theorem 1 holds.

(ii) Our argument can easily be applied to the problem

$$\begin{cases} -\Delta u + f(x,y,u) = 0 & \text{in } \Omega \\ u = g & \text{on } \Gamma \end{cases}$$

where $f \in C^2(\bar{\Omega} \times \mathbf{R})$ and $\frac{\partial f}{\partial u} \geq 0$ in $\bar{\Omega} \times \mathbf{R}$. This, together with the case where f is not necessarily smooth, will be disussed elsewhere.

Note : The content of this paper is a summany of an invited talk entitled "Revisit to finite difference methods in a bounded Dirichlet domain" by the author in the meeting "Study of Numerical Algorithms" organized by Prof. M.Mori which was held at RIMS, Kyoto University in November 27, 1997. The detail of the arguments including proofs of theorems and results of numerical experiments will be given in the forthcoming paper [7].

References

1. Y.-G. Chen, Blow-up solutions to a finite difference analogue of $u_t = \Delta u + u^{1+\alpha}$ in N-dimensional balls, Hokkaido Math. J. 21(1992), 447-474.
2. R.Courant and D.Hilbert, Methods of Mathematical Physics, vol. 2, Interscience Pub. 1962.
3. G.E.Forsythe and W.R.Wasaw : Finite Difference Methods for Partial Differential Equations, Springer 1977.
4. D.Gilbarg and N.S.Trudinger, Elliptic Partial Differential Equations of Second Order, Springer 1977
5. W.Hackbush, Elliptic Differential Equations, Springer 1992.
6. J.M.Ortega, Numerical Analysis, A Second Course, Academic Press 1972.
7. N.Matsunaga and T.Yamamoto, Superconvergence of finite difference methods for Dirichlet problems, in preparation.
8. A.A.Samarsky and V.B.Andreev, Difference Methods for Elliptic Equations (Russian), Nauka 1976.
9. J.C.Strikwerda, Finite Difference Schemes and Partial Difference Equations, Wadsworth & Brooks 1989.
10. J.C.Strikwerda and Y.Nagel, Finite difference methods for polar coordinate systems, MRC Technical Summary Report #2934, University of Wisconsin-Madison 1986.

11. P.N.Swartztrauber and R.A.Sweet, The direct solution of the discrete Poisson equation on a disk, SIAM J. Numer. Anal. 10(1973), 900-907.

12. R.S.Varga, Matrix Iterative Analysis, Prentice-Hall 1962.

13. D.M.Young, Iterative Solution of Large Linear Systems, Academic Press 1971.