

# Highly Accurate Numerical Solution of Casilinear Elliptic Boundary-Value Problems in $n$ Dimensions

By Victor Pereyra\*

**Abstract.** The method of Iterated Deferred Corrections, whose theory was developed by the author, is applied to the problems of the title. The necessary asymptotic expansions are obtained and the way in which the corrections are produced by means of numerical differentiation is described in detail. Numerical results and comparisons with the variational-splines methods are given.

**1. Introduction.** In several papers [6], [8], [9] the author has developed the theory of the method of Iterated Deferred Corrections (IDC), first introduced and extensively used by L. Fox. The most general results apply to nonlinear operator equations in Banach spaces and their abstract discretizations. Several important applications have been treated in detail [3], [7], [8], [9].

In the present publication we shall study the application of IDC to elliptic boundary-value problems in  $n$  dimensions. Previous work in partial differential equations has been restricted generally to the two-dimensional Poisson equation (see [8] for an extensive list of references and historical developments). In [8], [9] we described in detail the application of a linearized one-correction procedure to mildly nonlinear equations of the form  $\Delta u = f(x, y, u)$ , and gave numerical results.

In the first sections of this paper we treat the first boundary-value problem for a general linear uniformly elliptic operator with variable coefficients in  $n$  independent variables. This is extended to a casilinear equation, by which we mean an equation with a source term nonlinear both in the unknown function and its gradient.

Asymptotic expansions for the local discretization error are derived and suitable conditions are imposed in order that the IDC procedure be rigorously applicable. We use as the basic discretization one due to Pucci [10].

Finally in Section 7 we present some numerical results and comparisons with the high-order methods of Ciarlet et al. [2], as implemented by Herbold [5]. The conclusion of this fairly restricted comparison is that IDC is, in this case, about 100 times faster than the best variational-spline method. Of course, our method only gives results at the grid points; however, it is clear that once an accurate solution is obtained on some finite grid we can interpolate, or even produce multivariate splines in order to have an overall defined approximate solution, if that is desired.

**2. Preliminaries.** Let  $A$  be an open, bounded set in  $E^n$ , the  $n$ -dimensional Euclidean space. Let  $a_{i,j}(y) \equiv a_{i,j}(y)$ ,  $a_i(y)$  ( $i, j = 1, \dots, n$ ),  $a(y)$ ,  $f(y)$ , be  $C^M$  functions

---

Received September 29, 1969, revised March 17, 1970.

AMS 1968 subject classifications. Primary 65G5, 65G6.

Key words and phrases. Casilinear elliptic equations, boundary-value problems, deferred corrections, finite differences, high-order discrete methods.

\* This work was completed while the author was visiting the Mathematics Research Center, University of Wisconsin during the summer of 1969.

( $M > 0$ ) defined on  $\bar{A}$ ; assume that  $a \leq 0$ , and that for any vector  $\xi = (\xi_1, \dots, \xi_n)$

$$\sum_{i,j} a_{ij}(y)\xi_i\xi_j \geq \alpha|\xi|^2 \quad \text{on } A,$$

with  $\alpha > 0$ . Under these conditions the linear differential operator

$$(2.1) \quad Lu(y) = \sum_{i,j} a_{ij}(y) \frac{\partial^2 u(y)}{\partial y_i \partial y_j} + \sum_i a_i(y) \frac{\partial u(y)}{\partial y_i} + a(y)u(y)$$

is uniformly elliptic.

For  $\phi(y) \in C^{M+2}(\bar{A})$  we shall assume that the Dirichlet problem

$$(2.2) \quad \begin{aligned} Lu(y) &= f(y) \quad \text{on } A, \\ u(y) &= \phi(y) \quad \text{on } \partial A, \end{aligned}$$

has a solution  $u(y) \in C^{M+2}(\bar{A})$ .

In what follows we shall use interchangeably:

$$u_i \equiv \frac{\partial u}{\partial y_i}, \quad u_{ii} \equiv \frac{\partial^2 u}{\partial y_i \partial y_i}, \quad D_i^k u \equiv \frac{\partial^k u}{\partial y_i^k}.$$

We are interested in discretizing problem (2.2). For this we introduce a uniform mesh  $\Omega_h$  on  $E^n$ :

$$(2.3) \quad \Omega_h = \{x \in E^n: x = (r_1 h, r_2 h, \dots, r_n h), r_1, \dots, r_n \text{ arbitrary integers, } h > 0\},$$

and also  $A_h = A \cap \Omega_h, \bar{A}_h = \bar{A} \cap \Omega_h, \partial A_h = \partial A \cap \Omega_h$ .

All the results are valid, with minor changes, if we use instead a mesh with different step sizes in each coordinate direction.

Let  $h_i$  be a vector in the direction of the positive  $y_i$  axis and of modulus  $h$ . A point  $x \in A_h$  is called regular iff  $x \pm h_i \pm h_j \in \bar{A}_h$  ( $i, j = 1, \dots, n; i \neq j$ ). From now on we shall assume that *the region  $A$  and the mesh  $\Omega_h$  are such that  $A_h$  consists only of regular points.*

We define now some difference operators:

$$(2.4) \quad \begin{aligned} \Delta_i^+ u(x) &= \frac{u(x + h_i) - u(x)}{h}, & \Delta_i^- u(x) &= \frac{u(x) - u(x - h_i)}{h}, \\ \delta_i &= \frac{1}{2}(\Delta_i^+ + \Delta_i^-), & \Delta_{ii} &= \Delta_i^+ \Delta_i^-, \\ \Delta_{ij}^- &= \frac{1}{2}(\Delta_i^+ \Delta_j^- + \Delta_j^- \Delta_i^+), & \Delta_{ij}^+ &= \frac{1}{2}(\Delta_i^+ \Delta_j^+ + \Delta_j^+ \Delta_i^+). \end{aligned}$$

It is easy to verify that

$$\Delta_i^+ \Delta_j^- = \Delta_j^- \Delta_i^+, \quad \Delta_{ij}^+ = \Delta_{ji}^+, \quad \Delta_{ij}^- = \Delta_{ji}^-, \quad \Delta_{ii} = \Delta_i^- \Delta_i^+.$$

**3. The Discretization.** We discretize (2.2) in the following form [10]:

$$(3.1) \quad L_h u(y) = \sum_{i,j} a_{ij}(y) \Delta_{ij} u(y) + \sum_i a_i(y) \delta_i u(y) + a(y)u(y) = f(y), \quad y \in A_h,$$

where  $\Delta_{ij}$  ( $i \neq j$ ) is defined as:

$$(3.2) \quad \begin{aligned} \Delta_{ij} u(y) &= \Delta_{ij}^+ u(y) \quad \text{if } a_{ij}(y) \geq 0, \\ &= \Delta_{ij}^- u(y) \quad \text{if } a_{ij}(y) < 0. \end{aligned}$$

For the boundary points we have:

$$L_h u(y) \equiv u(y) = \phi(y), \quad y \in \partial A_h.$$

If we further assume that

$$(3.3) \quad 2a_{ii} - \sum_{i \neq j} |a_{ij}| > \bar{\alpha} > 0 \quad \text{on } \bar{A},$$

then for small  $h$ 's a maximum principle is valid for the operator  $L_h$ . That is, if  $B \geq 0$  is such that  $|a_i| \leq B$  on  $\bar{A}$ , and  $Bh < \bar{\alpha}$ , then  $f \leq 0$  implies that a solution  $u$  of

$$(3.4) \quad L_h u = f \quad \text{on } A_h, \quad u = \phi \quad \text{on } \partial A_h,$$

must satisfy:

$$u(x) \geq \max \left\{ 0, \min_{y \in \partial A_h} \phi(y) \right\}, \quad x \in A_h.$$

As usual, the maximum principle immediately implies that there exists a unique solution of the discrete problem (3.4), and also that an a priori bound for this solution can be obtained for  $x \in \bar{A}_h$ :

$$(3.5) \quad |u(x)| \leq \max_{x \in \partial A_h} |\phi(x)| + K \max_{x \in \bar{A}_h} |f(x)|,$$

where the positive constant  $K$  depends only on  $B$  and the domain  $A$ . For a proof of these statements we refer to [10].

The estimate (3.5) is in fact valid for any mesh function  $u(x)$ :

$$(3.6) \quad \|u(x)\|_\infty \leq K \|L_h u\|_\infty,$$

and therefore  $L_h$  is uniformly stable in the sense of [6, p. 317].

**4. Asymptotic Expansions.** Taking advantage of the differentiability assumptions we obtain now an asymptotic expansion in powers of  $h$  for  $L_h u(y)$ ,  $u \in C^{M+2}(\bar{A})$ . This will be valid, in particular, for the solution of (2.2). The expansion is obtained, as usual, by using a multiple Taylor formula with  $M + 1$  terms around  $y$ , for each one of the function values that appear in  $L_h u(y)$ . The remainder term will be of order  $M$  in  $h$ .

First of all we have:

$$(4.1) \quad \begin{aligned} & h^2 \sum_{i,j} a_{ij}(y) \Delta_{ij} u(y) \\ &= \sum_i \left\{ \left[ \sum_j |a_{ij}(y)| - 3a_{ii}(y) \right] u(y) \right. \\ & \quad + \left( 2a_{ii}(y) - \sum_j |a_{ij}(y)| \right) [u(y + h_i) + u(y - h_i)] \\ & \quad \left. + \frac{1}{2} \sum_{j: j \neq i} |a_{ij}(y)| [u(y + h_i + \epsilon_{ij} h_j) + u(y - h_i - \epsilon_{ij} h_j)] \right\}, \end{aligned}$$

where  $\epsilon_{ij} = \text{sign}(a_{ij}(y))$ , ( $\text{sign } 0 = 1$ ).

Expansions for the relevant terms are:

$$u(y + h_i) + u(y - h_i) = 2 \sum_{\nu=0}^N D_i^{2\nu} u(y) \frac{h^{2\nu}}{(2\nu)!} + O(h^{M+2}),$$

$$\begin{aligned}
 & u(y + h_i + \epsilon_{i,j}h_j) + u(y - h_i - \epsilon_{i,j}h_j) \\
 &= 2 \sum_{\nu=0}^N \left( \sum_{\mu=0}^{2\nu} \binom{2\nu}{\mu} \epsilon_{i,j}^\mu D_i^{2\nu-\mu} D_j^\mu u(y) \right) \frac{h^{2\nu}}{(2\nu)!} + O(h^{M+2}),
 \end{aligned}$$

where  $N = \lceil \frac{1}{2}(M + 1) \rceil$ .

Therefore,

$$\begin{aligned}
 h^2 \sum_{i,j} a_{i,j}(y) \Delta_{i,j} u(y) &= h^2 \sum_{i,j} a_{i,j}(y) u_{i,j}(y) \\
 (4.2) \quad &+ \sum_{\nu=2}^N \left\{ \sum_i \left[ 2a_{i,i}(y) D_i^{2\nu} u(y) \right. \right. \\
 &\quad \left. \left. + \sum_{j \neq i} |a_{i,j}(y)| \sum_{\mu=1}^{2\nu-1} \binom{2\nu}{\mu} \epsilon_{i,j}^\mu D_i^{2\nu-\mu} D_j^\mu u(y) \right] \right\} \frac{h^{2\nu}}{(2\nu)!} + O(h^{M+2}).
 \end{aligned}$$

Since

$$h\delta_i u(y) = hu_{i,j}(y) + \sum_{\nu=2}^N D_i^{2\nu-1} u(y) \frac{h^{2\nu-1}}{(2\nu-1)!} + O(h^{2N+1}),$$

we finally get

$$\begin{aligned}
 h^2 L_h u(y) &= h^2 Lu(y) \\
 &+ \sum_{\nu=2}^N \sum_i \left[ 2a_{i,i}(y) D_i^{2\nu} u(y) \right. \\
 (4.3) \quad &\quad \left. \cdot \sum_{j \neq i} |a_{i,j}(y)| \sum_{\mu=1}^{2\nu-1} \binom{2\nu}{\mu} \epsilon_{i,j}^\mu D_i^{2\nu-\mu} D_j^\mu u(y) \right. \\
 &\quad \left. + (2\nu)a_{i,i}(y) D_i^{2\nu-1} u(y) \right] \frac{h^{2\nu}}{(2\nu)!} \\
 &+ O(h^{M+2}), \quad y \in A_h,
 \end{aligned}$$

where the terms in the  $\nu$ -sum are a detailed expression for the local truncation error  $L_h u(y) - Lu(y)$ .

By taking  $N = 2$  we obtain

$$(4.4) \quad L_h u(y) = Lu(y) + O(h^2),$$

i.e., the discrete method is consistent of order two.

By using (3.6) and (4.4) on the mesh function  $e(y) = v(y) - u(y)$  where  $v(y)$  satisfies

$$L_h v(y) = f(y), \quad y \in A_h, \quad v(y) = \phi(y), \quad y \in \partial A_h,$$

and  $u(y)$  satisfies  $Lu = f, u = \phi$ , we see that:

$$\|e(y)\| \leq K \|L_h e\| = O(h^2),$$

and the discrete method is convergent of order  $h^2$ .

We finally assume that the global discretization error  $e(y)$  satisfies

$$(4.5) \quad e(y) = \sum_{\nu=1}^{N-1} e_\nu(y) h^{2\nu} + O(h^M)$$

where the  $e_\nu(y) \in C^{M-2\nu+2}(\bar{A})$  are independent of  $h$ .

This is, of course, the type of asymptotic expansion that makes also possible the use of successive extrapolations to the limit [8].

**5. Iterated Deferred Corrections (IDC).** The results and hypotheses of the earlier sections allow the application of deferred corrections in a rigorous manner, in order to obtain a stable discrete method of high asymptotic order in  $h$ .

The crucial point now is the construction of sufficiently accurate formulae to approximate the segments of the local truncation error (4.3).

If all the  $a_{i,i}(y)$  ( $i \neq j$ ) are identically zero, then the problem takes the form:

$$(5.1) \quad Mu(y) \equiv \sum_{i=1}^n \left[ a_{i,i}(y) \frac{\partial^2 u(y)}{\partial y_i^2} + a_i(y) \frac{\partial u(y)}{\partial y_i} \right] + a(y)u(y) = f(y) \quad \text{on } A,$$

$$u(y) = \phi(y) \quad \text{on } \partial A.$$

This simplified form, with no cross derivatives, can be dealt with as a Cartesian product of one-dimensional problems, i.e. each coordinate direction (or grid line) can be treated independently, using one-dimensional techniques.

The general case (2.1), with the discretization (3.1), does require truly multi-dimensional techniques which, due to the order of the equation being considered, have only to be two-dimensional since only derivatives in at most two directions are present at any time. As this poses new problems, we shall delay its discussion to a later paper. At this time we shall only say that fast and accurate techniques for multidimensional numerical differentiation are available [4]; they have been tested for two and three dimensions and will soon be applied to problem (2.1).

We concentrate our efforts now on the simplified form (5.1) and its discretization:

$$(5.2) \quad M_h u(y) \equiv \sum_{i=1}^n [a_{i,i}(y)\Delta_{i,i}u(y) + a_i(y)\delta_i u(y)] + a(y)u(y) = f(y), \quad y \in A_h,$$

$$M_h u(y) \equiv u(y) = \phi(y), \quad y \in \partial A_h.$$

Observe that now the requirement for a point to be regular is less strict:  $y \in A_h$  is a regular point iff  $y \pm h_i \in \bar{A}_h$ .

The asymptotic expansion for the local discretization error reduces to:

$$(5.3) \quad h^2 M_h u(y) = h^2 Mu(y) + \sum_{i=1}^n \sum_{\nu=2}^N \{2a_{i,i}(y)D_i^{2\nu}u(y) + 2\nu a_i(y)D_i^{2\nu-1}u(y)\} \frac{h^{2\nu}}{(2\nu)!}$$

$$+ O(h^{M+2}), \quad y \in A_h.$$

We further require in what follows that the region  $A$  be convex. This is by no means indispensable but makes the ensuing developments simpler.

Let us consider now an arbitrary grid line, say in the direction of the  $i$ th coordinate. All coordinates but the  $i$ th one are fixed, say to the values  $\bar{y}_1, \dots, y_{i-1}, y_{i+1}, \dots, \bar{y}_n$ . Let  $\bar{y}$  be the vector that describes the coordinates of a point moving on that line, and let  $\bar{y}_{LB}, \bar{y}_{RB}$  be the coordinates of the intersections of the grid line with the boundary  $\partial A$ . If we consider the one-dimensional boundary value problem

$$a_{i,i}(\bar{y}) \frac{\partial^2 u(\bar{y})}{\partial y_i^2} + a_i(\bar{y}) \frac{\partial u(\bar{y})}{\partial y_i} = 0,$$

$$u(\bar{y}_{LB}) = \phi(\bar{y}_{LB}), \quad u(\bar{y}_{RB}) = \phi(\bar{y}_{RB}),$$

and its corresponding discretization

$$a_{ii}(\bar{y})\Delta_{ii}u(\bar{y}) + a_i(\bar{y})\delta_i u(\bar{y}) = 0,$$

$$u(\bar{y}_{LB}) = \phi(\bar{y}_{LB}), \quad u(\bar{y}_{RB}) = \phi(\bar{y}_{RB}),$$

then the inner sum in (5.3) is exactly the representation of the local truncation error for this problem (setting  $y \equiv \bar{y}$ ). This allows us to use the same technique we employed in [7, Sections 2, 3] in order to generate approximations to the segments of the local truncation error. At each point  $y$  we have just to produce appropriate *one-dimensional* differentiation formulae in each coordinate direction, and then add them all up in order to obtain:

$$(5.4) \quad S_k u(y) = \sum_{i=1}^n \sum_{\nu=1}^k \{2a_{ii}(y)D_i^{2\nu+2}u(y) + (2\nu + 2)a_i(y)D_i^{2\nu+1}u(y)\} \frac{h^{2\nu+2}}{(2\nu + 2)!} + O(h^{2k+4}), \quad k = 1, \dots, N - 1.$$

Just for the sake of completeness we recall that the one-dimensional algorithm consisted essentially of the automatic generation of numerical differentiation formulae using linear combinations of ordinates. An improved version of the algorithm used in [7] can be found in [1].

In the present case we shall have formulae like:

$$(5.5) \quad S_k u(y) = \sum_{i=1}^n \sum_{j=1}^{J_i} \omega_{ij}(y)u(y^{ij})$$

where  $J_i = 2k + 4$  or  $2k + 3$  according to the position of the point  $y$  on the grid line passing through it and parallel to the  $i$ th direction. The value  $2k + 3$  is used for those points that have at least  $k$  regular neighbors at each side of them on the grid line mentioned above. Counting also the boundary intersections we can center on  $y$  a  $2k + 3$  points symmetric formula in the direction  $i$ , which has the correct order of accuracy. If it is not possible to center such a formula on  $y$ , for instance when we approach the boundary, then a  $2k + 4$  points unsymmetric formula is necessary. We always use the "most centered formula" possible since that decreases the truncation error and the size of the weights as well.

Once we have assured the construction of  $S_k U$ , for any grid function  $U$ , then we can describe the deferred corrections procedure.

- (1) Solve (5.2) in order to obtain  $U^{(0)}(y)$ ,  $y \in A_h$ .
- (2) Compute  $S_1 U^{(0)}(y)$ ,  $y \in A_h$ .
- (3) Solve

$$(5.6)_0 \quad h^2 M_h \Delta U(y) = S_1 U^{(0)}(y), \quad y \in A_h,$$

$$\Delta U(y) = 0, \quad y \in \partial A_h.$$

- (4) Correct  $U^{(1)}(y) = U^{(0)}(y) + \Delta U(y)$  ( $\Delta U$  means increment here, not Laplacian). Iterate.

In the  $k$ th step we shall have:

$$(5.6)_k \quad h^2 M_h \Delta U(y) = S_k(U^{(k-1)}(y)) - S_{k-1}(U^{(k-2)}(y)), \quad y \in A_h,$$

$$\Delta U(y) = 0, \quad y \in \partial A_h,$$

$$U^{(k)}(y) = U^{(k-1)}(y) + \Delta U(y), \quad \|U^{(k)}(y) - u(y)\|_\infty = O(h^{2k+2}).$$

If we are worried about storage, we can save some by solving directly for a corrected value rather than for a correction. The interest of correcting as indicated in formula (5.6) is that the  $\Delta U(y)$  of the  $k$ th step is an asymptotic error estimate for the error in the  $(k - 1)$ th step. This statement is true in the general nonlinear operator case and we append a proof at the end of this paper.

We shall prove here the statement for problem (5.1). If  $u(y)$  is the solution of (5.1) and we assume that the  $S_k$  ( $k = 1, 2, \dots$ ) satisfying (5.4) can be constructed, then we have

(i)  $h^2(M_h u(y) - f(y)) = F_k(y) + O(h^{2k+4})$ , where  $F_k(y)$  represents the sum of the first  $k$  terms of the local truncation error (see (5.3) and (5.4)).

(ii) From (5.4) and Lemma 3.1 of [6]:

$$S_k U^{(k-1)}(y) = F_k(y) + O(h^{2k+4}).$$

Since  $U^{(k-1)}$  satisfies

$$h^2 M_h U^{(k-1)}(y) = h^2 f(y) + S_{k-1} U^{(k-2)}(y),$$

we have

$$h^2 M_h [u(y) - U^{(k-1)}(y)] = F_k(y) - S_{k-1} U^{(k-2)}(y) + O(h^{2k+4}),$$

and by (ii):

$$h^2 M_h [u(y) - U^{(k-1)}(y)] = S_k U^{(k-1)}(y) - S_{k-1} U^{(k-2)}(y) + O(h^{2k+4}).$$

From this last expression and the stability of the operator  $M_h$  ((3.6)) it follows that:

$$\|u(y) - U^{(k-1)}(y)\| = \|(h^2 M_h)^{-1} [S_k U^{(k-1)}(y) - S_{k-1} U^{(k-2)}(y)]\| + O(h^{2k+2}).$$

Since we already know that  $\|u(y) - U^{(k-1)}(y)\|$  is  $O(h^{2k})$ , the solution  $\Delta U(y)$  of (5.6) gives the most significant part of the error.

**6. The Casilinear Case.** We extend now the results of Section 5 to the casilinear case, by which we mean a problem of the form:

$$(6.1) \quad \begin{aligned} Lu(y) &= \sum_{i,j} a_{ij}(y) \frac{\partial^2 u(y)}{\partial y_i \partial y_j} + f(y, u, \nabla u) = 0, & y \in A, \\ u(y) &= \phi(y), & y \in \partial A, \end{aligned}$$

where the second order part of  $L$  is a uniformly elliptic operator and satisfies (3.3). The difference here is, of course, that  $u$  and  $\nabla u$  appear nonlinearly on the source term. If we further assume that for  $y \in A$ , all  $u, z$ , and a positive constant  $B$ :

$$(6.1') \quad f_u(y, u, z) \leq 0 \quad \text{and} \quad |f_{z_i}(y, u, z)| \leq B,$$

then an argument similar to that of Bers [11] assures the existence and uniqueness of a solution of problem (6.1).

If we discretize  $\nabla u(y)$  by means of

$$\nabla_h u(y) = (\delta_1 u(y), \dots, \delta_n u(y))$$

then we can define

$$(6.2) \quad \begin{aligned} L_h u(y) &= \sum_{i,j} a_{ij}(y) \Delta_{ij} u(y) + f(y, u, \nabla_h u) = 0, & y \in A_h, \\ u(y) &= \phi(y), & y \in \partial A_h. \end{aligned}$$

The linearized form of  $L_h u(y)$  is

$$(6.3) \quad L'_h[u(y)]e(y) = \sum_{i,j} a_{i,j}(y)\Delta_{i,j}e(y) + \sum_i fp_i(y, u, \nabla_h u)\delta_i e(y) + f_u(y, u, \nabla_h u)e(y) = 0,$$

where  $p_i = \partial u / \partial y_i$ .

If we compare (6.3) with (3.1) we see that for a given  $u$ ,  $fp_i(y, u(y), \nabla_h u(y))$  and  $f_u(y, u(y), \nabla_h u(y))$  take the place of  $a_i(y)$  and  $a(y)$  respectively.

If  $\bar{\alpha}$  is as in (3.3) the constant corresponding to the linear part of  $Lu(y)$ , then  $Bh < \bar{\alpha}$  assures the maximum principle for  $L'_h[u(y)]$  (uniformly in  $u$ ).

In this case the discretization  $L_h u$  is uniformly stable and an application of Theorem 2.1 of [6] gives us discrete convergence of order 2, as in the linear case.

Observe that the presence of derivatives in the nonlinear part makes the discretization itself nonlinear. This presents some difficulties in the generation of the asymptotic expansions, and in particular we have to require that  $f(y, u, z)$  be  $M$  times continuously differentiable in all its arguments. With this assumption we can proceed to analyze the nonlinear part in the following form. Firstly we expand

$$(6.4) \quad f(y, u, \nabla_h u) = f\left(y, u, \nabla u + \sum_{\nu=1}^{N-1} D^{2\nu+1}u(y) \frac{h^{2\nu}}{(2\nu+1)!} + O(h^{2N})\right)$$

in a multiple Taylor formula around  $(y, u, \nabla u)$ . Here

$$D^{2\nu+1}u(y) \equiv \left(\frac{\partial^{2\nu+1}u(y)}{\partial y_1^{2\nu+1}}, \dots, \frac{\partial^{2\nu+1}u(y)}{\partial y_n^{2\nu+1}}\right).$$

If we put

$$\delta = (\delta_i)_{i=1, \dots, n} \equiv \sum_{\nu=1}^{N-1} D^{2\nu+1}u(y) \frac{h^{2\nu}}{(2\nu+1)!},$$

and

$$\langle \nabla_p \cdot, \delta \rangle f(y, u, \nabla u) \equiv \sum_{i=1}^n \frac{\partial f(y, u, \nabla u)}{\partial p_i} \cdot \delta_i,$$

then the desired Taylor expansion is

$$(6.5) \quad f(y, u, \nabla_h u) = f(y, u, \nabla u) + \sum_{\mu=1}^{N-1} \frac{1}{\mu!} \langle \nabla_p \cdot, \delta \rangle^\mu f(y, u, \nabla u) + O(h^{2N}),$$

where the powers  $\langle \nabla_p \cdot, \delta \rangle^\mu$  are to be interpreted as operator compositions. The problem is that (6.5) is not an explicit expansion in powers of  $h$  yet. In order to proceed further we introduce the following standard notation:

$\sigma \equiv (\sigma_i)_{i=1, \dots, n}$ ,  $\sigma_i$  nonnegative integers;

$$|\sigma| \equiv \sum_{i=1}^n \sigma_i; \quad \delta^\sigma \equiv \prod_{i=1}^n \delta_i^{\sigma_i}; \quad \sigma! = \prod_{i=1}^n \sigma_i!$$

We first observe that

$$(6.5') \quad \begin{aligned} \frac{1}{\mu!} \langle \nabla_p \cdot, \delta \rangle^\mu f(y, u, \nabla u) &= \sum_{|\sigma|=\mu} \frac{\partial^{|\sigma|} f(y, u, p_1, \dots, p_n)}{\partial p_1^{\sigma_1} \dots \partial p_n^{\sigma_n}} \frac{\delta^\sigma}{\sigma!} \\ &= \sum_{|\sigma|=\mu} \nabla_p^\sigma f(y, u, \nabla u) \frac{\delta^\sigma}{\sigma!}. \end{aligned}$$

An algebraic manipulation gives

$$(6.6) \quad \delta^\sigma = \sum_{\nu=|\sigma|}^{N-1} \alpha_{\sigma\nu}[u(y)]h^{2\nu} + O(h^{2N}),$$

where the coefficients  $\alpha_{\sigma\nu}[u(y)]$ , obtained by collecting terms in equal powers of  $h$ , involve products of  $D_i^\nu u(y)$  and numerical constants, but no  $h$ 's.

Now we obtain from (6.5), (6.5'), and (6.6):

$$f(y, u, \nabla_h u) = f(y, u, \nabla u) + \sum_{\mu=1}^{N-1} \sum_{|\sigma|=\mu} \frac{1}{\sigma!} \nabla_p^\sigma f(y, u, \nabla u) \sum_{\nu=\mu}^{N-1} \alpha_{\sigma\nu}[u(y)]h^{2\nu} + O(h^{2N}).$$

Or

$$(6.7) \quad f(y, u, \nabla_h u) = f(y, u, \nabla u) + \sum_{\nu=1}^{N-1} \left\{ \sum_{\mu=1}^{\nu} \sum_{|\sigma|=\mu} \frac{1}{\sigma!} \nabla_p^\sigma f(y, u, \nabla u) \alpha_{\sigma\nu}[u(y)] \right\} h^{2\nu} + O(h^{2N}).$$

Recalling from Section 4 the asymptotic expansion for the discretization of the linear part of operator (6.1), and adding to it the corresponding one for the non-linear part just derived, we finally obtain the asymptotic expansion for the local discretization error in the casilinear case:

$$(6.7') \quad \begin{aligned} h^2 L_h u(y) &= h^2 L u(y) \\ &+ \sum_{\nu=2}^N \left\{ \sum_{i=1}^n \left[ 2a_{ii}(y) D_i^{2\nu} u(y) + \sum_{j \neq i} |a_{ij}(y)| \sum_{\mu=1}^{2\nu-1} \binom{2\nu}{\mu} \epsilon_{ij}^\mu D_i^{2\nu-\mu} D_j^\mu u(y) \right] \frac{1}{(2\nu)!} \right. \\ &\quad \left. + \sum_{\mu=1}^{\nu-1} \sum_{\|\sigma\|=\mu} \frac{1}{\sigma!} \nabla_p^\sigma f(y, u, \nabla u) \alpha_{\sigma, \nu-1}[u(y)] \right\} h^{2\nu} \\ &+ O(h^{M+2}), \quad y \in A_h. \end{aligned}$$

This asymptotic expansion in even powers of  $h$ , together with the general results of [6, Lemma 3.1 and Theorem 3.2] allow us, theoretically at least, to proceed with the IDC method in this nonlinear case. However, it is clear from (6.7') that the general application of IDC may present insurmountable difficulties.

There are two simplified cases of interest: (a) when  $f$  does not depend upon  $\nabla u$ ; (b) when  $\nabla u$  enters only linearly in  $f$ , i.e.  $f$  is of the form

$$\begin{aligned} f(y, u, \nabla u) &\equiv g(y, u) + \sum_{i=1}^n a_i(y, u) \frac{\partial u}{\partial y_i} \\ &\equiv g(y, u) + \langle a(y, u), \nabla u \rangle. \end{aligned}$$

For (a) no discretization of the nonlinear part is needed, thus the expansion is the same as in the linear case. For (b) we have that  $\nabla_p^\sigma f \equiv 0$  for  $|\sigma| > 1$ , and therefore

$$(6.8) \quad \begin{aligned} f(y, u, \nabla_h u) &= g(y, u) + \langle a(y, u), \nabla_h u \rangle \\ &= g(y, u) + \langle a(y, u), \nabla u \rangle \\ &\quad + \sum_{\nu=1}^{N-1} \langle a(y, u), D^{2\nu+1} u(y) \rangle \frac{h^{2\nu}}{(2\nu+1)!} + O(h^{2N}). \end{aligned}$$

In this case the nonlinear part of the local truncation error<sup>m</sup> to be approximated is fairly simple, and at the step  $k$  of our correction process we can replace

$$\Gamma_k u(y) \equiv \sum_{\nu=1}^k D^{2\nu+1} u(y) \frac{h^{2\nu}}{(2\nu+1)!}$$

by  $T_k$ , an  $O(h^{2k+2})$  approximation using only ordinate values of  $u(y)$  ( $y \in \bar{A}_h$ ).

Since for any smooth function  $u$ ,

$$\langle a(y, u), \Gamma_k u(y) \rangle = \langle a(y, u), T_k u(y) \rangle + O(h^{2k+2}), \quad y \in A_h,$$

then it follows from Lemma 3.1 of [6] that  $T_k U^{(k-1)}$  will be adequate for use in the IDC procedure.

**7. Numerical Results.** We give now some numerical results obtained with a computer implementation of the IDC procedure. First we present results for a linear problem:

$$\begin{aligned} \Delta u &= -2 \sin(x + y), & (x, y) \in D, \\ u &= \sin(x + y), & (x, y) \in \partial D, \end{aligned}$$

where  $D$  is the square of side  $l$ . The exact solution of this problem is  $u(x, y) = \sin(x + y)$ .

The basic discretization is the standard five point formula for the Laplacian. We used point SOR with parameter  $\omega$  for solving the linear equations at every correction step. The results were obtained on the CDC 3600 computer at the University of Wisconsin Computing Center, with a FORTRAN 63 program operating in double precision mode (-25 decimal digits).

For  $l = 1$ ,  $h = 1/16$ ,  $\omega = 1.6735$ , and  $k$  the correction number we obtained the results shown in Table I. The column headed  $\epsilon_k$  gives the maximum (exact) relative error at the grid points, while  $EST_k$  gives the value estimated by using the procedure of Section 5. From Section 5 we have that

$$EST_k = \max_{y \in A_h} |U^{(k+1)}(y) - U^{(k)}(y)| / |U^{(k+1)}(y)|.$$

By using the values  $EST_k$  we have implemented an automatic stopping criteria

TABLE I

$k$	$\epsilon_k$	$EST_k$	SOR iterat.	max. residual
0	4.6 (-5)	4.7 (-5)	41	9.0 (-7)
1	1.3 (-7)	1.3 (-7)	68	8.0 (-15)
2	5.7 (-9)	2.6 (-8)	67	3.9 (-17)
3	2.4 (-8)	7.1 (-8)	73	1.4 (-19)

Automatic stop since  $EST_3 > EST_2$ .

for the sequence of corrections. The process is interrupted if at any time

$$\text{either } EST_k \geq EST_{k-1} \text{ or } EST_k \leq \epsilon,$$

where  $\epsilon$  is the maximum relative tolerance supplied by the user. This criteria is seen at work in the results of Table I, and it has been equally effective in other problems run by the author.

In Table II we give results for the same problem with  $l = 0.1, h = 1/160, 1/320$ . Here we have used a fixed number of terms ( $k = 4$ ) in the corrections and have solved the linear equations to full accuracy all the time (max. res.  $< 10^{-23}$ ). The  $j$  in the first column indicates the correction number.

Finally we consider a mildly nonlinear equation that has been solved in [5, p. 184] by the methods of Ciarlet et al. [2]. The problem is:

$$\begin{aligned} \Delta u = u^3 + (-2 + (1 - 2x)^2)(e^{y(1-y)} - 1 + u) + (-2 + (1 - 2y)^2) \\ \cdot (e^{x(1-x)} - 1 + u) - (e^{x(1-x)} - 1)^3(e^{y(1-y)} - 1)^3, \text{ in } D, \\ u(x, y) = 0, \text{ in } \partial D, \end{aligned}$$

where  $D$  is the unit square.

We use Newton's method and point SOR in order to solve the linear equations at every Newton step. An adequate strategy is employed in order to terminate the several nested iterations producing an overall efficient procedure [8].

We compare in Table III the two best methods of Herbold (in terms of computing time) with IDC. The result with only one correction was obtained as a part of the run with three corrections and therefore the time is only estimated. The time can be improved if we want to perform only one correction, since that can be obtained with a linearized procedure [8], [9].

We are neither claiming nor implying that the IDC procedure should be preferred to the variational-spline approach. The only point we would like to make is that in this type of problem the numerical results for IDC are quite encouraging.

**Appendix.** Using the notation of [6] we shall prove that a posteriori asymptotic error estimates can be obtained in the general nonlinear operator case. We recall some relevant facts:

- (i)  $\Phi_h(\varphi_h x^*) = \psi_h \hat{r}_{k+1}(x^*) + O(h^{p \cdot (k+2)});$
- (ii)  $S_{k+1}(\varphi_h x^*) - \psi_h \hat{r}_{k+1}(x^*) = O(h^{p \cdot (k+2)});$

TABLE II

$j$	$\epsilon(1/160)$	$\epsilon(1/320)$
0	5.0 (-9)	1.3 (-9)
1	1.5 (-11)	9.3 (-13)
2	5.6 (-13)	3.4 (-14)
3	1.1 (-13)	6.5 (-15)
4	2.7 (-14)	1.6 (-15)
5	8.0 (-15)	4.8 (-16)

TABLE III

Method	Maximum relative error	Time in seconds	Computer
Herbold-Ciarlet . . . Hermite $H_0^2(\pi_N)$	5.0 (-5)	3388	UNIVAC 1107†
Same Splines $S_{p_0}^2(\pi_N)$	9.0 (-5)	2450	UNIVAC 1107
Deferred Corrections 1 correct.	2.0 (-5)	200 (estimated)	IBM-360/40
Same 3 correct.	5.0 (-7)	327	IBM-360/40

† The UNIVAC 1107 is about ten times faster than the IBM-360/40. This computation was carried out at the Departamento de Computación, Facultad de Ciencias, Universidad Central de Venezuela, Caracas.

$$(iii) \quad S_{k+1}(\varphi_h x^*) - S_{k+1}(U^{(k)}) = O(h^{p \cdot (k+2)});$$

$$(iv) \quad \Phi_h(U^{(k)}) = S_k(U^{(k-1)}).$$

Combining (i) and (iv) we get

$$\Phi_h(\varphi_h x^*) - \Phi_h(U^{(k)}) = \psi_h \hat{I}_{k+1}(x^*) - S_k(U^{(k-1)}) + O(h^{p \cdot (k+2)}),$$

and by (ii), (iii) and the mean value theorem,

$$\Phi_h'(U^{(k)})(\varphi_h x^* - U^{(k)}) = S_{k+1}(U^{(k)}) - S_k(U^{(k-1)}) + O(h^{p \cdot (k+2)}).$$

From this last expression and the stability of  $\Phi_h$  we get

$$\|\varphi_h x^* - U^{(k)}\| = \|[\Phi_h'(U^{(k)})]^{-1}(S_{k+1}(U^{(k)}) - S_k(U^{(k-1)}))\| + O(h^{p \cdot (k+2)}).$$

The first part of the right-hand side can be computed by solving a linear problem after  $U^{(k)}$  has been obtained. Since it is already known that  $\|\varphi_h x^* - U^{(k)}\| = O(h^{p \cdot (k+1)})$ , this computation gives us the most significant part of the global discretization error.

**Acknowledgment.** I would like to thank Professor J. Barkley Rosser for his careful reading of the manuscript and for making valuable suggestions to improve its readability, especially where logical and notational coherence is concerned.

1. Å. BJÖRCK & V. PEREYRA, *Solution of Vandermonde Systems of Equations*, Pub. 70-02, Dept. de Comp., Fac. Ciencias, Univ. Central de Venezuela, Caracas, 1970; *Math. Comp.*, v. 24, 1970, pp. 893-903.
2. P. G. CIARLET, M. H. SCHULTZ & R. S. VARGA, "Numerical methods of high-order accuracy for nonlinear boundary value problems. V: Monotone operator theory," *Numer. Math.*, v. 13, 1969, pp. 51-77.
3. J. DANIEL, V. PEREYRA & L. SCHUMAKER, "Iterated deferred corrections for initial value problems," *Acta Ci. Venezolana*, v. 19, 1968, pp. 128-135.
4. G. GALIMBERTI & V. PEREYRA, "Numerical differentiation and the solution of multi-dimensional Vandermonde systems," Pub. 69-07, Dept. de Comp., Fac. Ciencias, Univ. Central de Venezuela, Caracas, 1969; *Math. Comp.*, v. 24, 1970, pp. 357-364.
5. R. J. HERBOLD, *Consistent Quadrature Schemes for the Numerical Solution of Boundary Value Problems by Variational Techniques*, Ph.D. Thesis, Case Western Reserve University, Cleveland, Ohio, 1968.
6. V. PEREYRA, "Iterated deferred corrections for nonlinear operator equations," *Numer. Math.*, v. 10, 1967, pp. 316-323. MR 36 #4812.
7. V. PEREYRA, "Iterated deferred corrections for nonlinear boundary value problems," *Numer. Math.*, v. 11, 1968, pp. 111-125. MR 37 #1091.
8. V. PEREYRA, "Accelerating the convergence of discretization algorithms," *SIAM J. Numer. Anal.*, v. 4, 1967, pp. 508-533. MR 36 #4778.
9. V. PEREYRA, "On improving an approximate solution of a functional equation by deferred corrections," *Numer. Math.*, v. 8, 1966, pp. 376-391. MR 34 #3814.
10. C. PUCCI, *Some Topics in Parabolic and Elliptic Equations*, Inst. for Fluid Dynamics, Lecture Series, 36, University of Maryland, College Park, Md., 1958.
11. L. BERS, "On mildly nonlinear partial differential equations of elliptic type," *J. Res. Nat. Bur. Standards*, v. 51, 1953, pp. 229-236. MR 16, 260.