

Purdue University
Purdue e-Pubs

Department of Computer Science Technical
Reports

Department of Computer Science

1984

The Importance of Scaling for the Hermite Bicubic Collocation Equations

Wayne R. Dyksen

John R. Rice
Purdue University, jrr@cs.purdue.edu

Report Number:
84-477

Dyksen, Wayne R. and Rice, John R., "The Importance of Scaling for the Hermite Bicubic Collocation Equations" (1984). *Department of Computer Science Technical Reports*. Paper 397.
<https://docs.lib.purdue.edu/cstech/397>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries.
Please contact epubs@purdue.edu for additional information.

THE IMPORTANCE OF SCALING FOR THE HERMITE BICUBIC
COLLOCATION EQUATIONS

Wayne R. Dyksen[†]

John R. Rice[†]

Department of Computer Sciences
Purdue University
West Lafayette, Indiana 47907

CSD-TR 477
March 1984
Revised August 1984

ABSTRACT

It is well known that improper scaling of linear equations can result in catastrophic loss of accuracy from Gauss elimination. The scaling process is not well understood and the commonly used "scaling rules" can fail. We study the scaling problem for the linear equations that arise from solving elliptic partial differential equations by collocation using Hermite bicubics. We present an *a priori* scaling rule that is effective but not foolproof. We conclude that one should use scaled partial pivoting for such equations. We also explore the relationship between the ordering used during Gauss elimination and the underlying geometry of the elliptic problem; we conjecture that this ordering must maintain the geometric integrity of the problem in order to avoid severe round-off problems.

[†]This work supported in part by Department of Energy contract DE-AC02-81ER10997

The Importance of Scaling for the Hermite Bicubic Collocation Equations

Wayne R. Dyksen
John R. Rice

1. Introduction

It is well known that improper scaling of linear equations often results in a catastrophic loss of accuracy from Gauss elimination. Unfortunately, the scaling process is not well understood and the most commonly used "scaling rules" can fail. Textbooks usually choose one of three courses: 1) say that the linear equations should be "properly scaled" and ignore the issue [Dongarra, et. al., 1979], 2) give some rules for scaling and then warn that they are not infallible [Rice, 1981], or 3) present scaled partial pivoting as the proper version of Gauss elimination [Conte and de Boor, 1980]. A few books combine these [Rice, 1983].

We report here on an experimental study of the scaling problem for the linear systems that arise from solving elliptic partial differential equations using Hermite bicubic collocation. An attractive feature of collocation is that it applies easily to general partial differential equations with general boundary conditions. However, the system of linear equations obtained from Hermite bicubic collocation does not possess any special properties such as being positive definite and, as a result, it is most often solved using simple band Gauss elimination. This study demonstrates that it is essential to scale the Hermite bicubic collocation equations; that is, if some type of scaling is not used, then the accumulated effects of round-off dominate the computations. We recommend using both a particular *a priori* scaling of the equations together with scaled partial pivoting. However, since we cannot formulate a completely reliable *a priori* scaling rule for these equations which requires less computation than what

scaling adds to scaled partial pivoting, we conclude that one should always use scaled partial pivoting. We believe that this conclusion is applicable to other finite element methods. Moreover, we conjecture that the ordering used during Gauss elimination must preserve the underlying geometry of an elliptic problem.

2. Collocation with Hermite Bicubics

We consider a second order, linear elliptic problem on a rectangular domain \mathbf{R} in the form

$$\begin{aligned} L[u] &= au_{xx} + bu_{xy} + cu_{yy} + du_x + eu_y + fu = g & (x, y) \in \mathbf{R} \\ M[u] &= \alpha u + \beta u_x + \gamma u_y = \delta & (x, y) \in \partial\mathbf{R} \end{aligned}$$

where $a, b, c, d, e, f, g, \alpha, \beta, \gamma$ and δ are given functions of x and y . We choose a positive integer n and subdivide the domain \mathbf{R} with a tensor product grid containing n^2 rectangles. We then approximate $u(x, y)$ by

$$U(x, y) = \sum_{i=1}^N H_i(x, y) \sim u(x, y)$$

where $N = 4(n+1)^2$ and the $H_i(x, y)$ are the Hermite bicubic basis functions formed as the tensor product of the standard one dimensional Hermite cubics with the grid lines being the knots.

The N unknowns w_i are determined by choosing N distinct points in \mathbf{R} and collocating the elliptic problem at these points. In particular, $4n^2$ collocation points are placed at the four Gauss points of each of the n^2 grid rectangles since this gives a fourth order discretization error for smooth problems [Houtis, 1978], [Percei and Wheeler, 1980]. The remaining $4(2n+1)$ collocation points are the two Gauss points of each boundary grid segment plus one at each of the four corners of \mathbf{R} . Collocating at these points, we obtain the *Hermite bicubic collocation equations*

$$L[U](x_k, y_k) = g(x_k, y_k) \quad k = 1, \dots, 4n^2$$

$$M[U](x_k, y_k) = \delta(x_k, y_k) \quad k = 4n^2 + 1, \dots, 4(n+1)^2.$$

The structure of the coefficient matrix of the resulting linear system is determined by the ordering of the collocation points (the equations or rows) and the basis functions (the unknowns or columns). A common finite element ordering is to order the grid rectangles in the natural way from bottom to top, left to right. The collocation points are then numbered corresponding to their containing grid rectangles (See Figure 4). The Hermite bicubic basis functions are ordered corresponding to their support in a natural way from bottom to top, left to right. The resulting coefficient matrix is somewhat block bi-diagonal [Dyksen, 1981], [Dyksen and Rice, 1982].

3. Numerical Experiment

We studied this problem using the ELLPACK system [Rice and Boisvert, 1984]. Its discretization modules P3C1 COLLOCATION and HERMITE COLLOCATION[†] generate the Hermite bicubic collocation equations; HERMITE COLLOCATION scales the equations associated with the boundary conditions (see Section 6) whereas P3C1 COLLOCATION does not. The solution modules LINPACK BAND and BAND GE solve the resulting linear system of equations. LINPACK BAND uses the LINPACK routines SGBFA and SGBSL which do band Gauss elimination with partial pivoting [Dongarra, et. al., 1979]. BAND GE does band Gauss elimination with scaled partial pivoting using a direct modification of SGBFA and SGBSL. The equations are solved in the order in which they are generated by the discretization modules, namely, the finite element ordering described above.

We combine these modules to obtain four similar, yet distinct numerical methods. Note that scaling is the only difference between P3C1 COLLOCATION and HERMITE COLLO-

[†]The module which we refer to here as HERMITE COLLOCATION has subsequently been split into two separate ELLPACK modules, HERMITE COLLOCATION and INTERIOR COLLOCATION, and P3C1 COLLOCATION has been removed from ELLPACK.

CATION and between LINPACK BAND and BAND GE. Letting S and U stand for "scaled" and "unscaled", we denote these four methods as follows:

Notation	Numerical Method
U/U	P3C1 COLLOCATION/LINPACK BAND
U/S	P3C1 COLLOCATION/BAND GE
S/U	HERMITE COLLOCATION/LINPACK BAND
S/S	HERMITE COLLOCATION/BAND GE

We use a subject population of twenty elliptic problems from the population of [Rice, et. al., 1981]; it consists of problems 2-1, 3-1, 5-1, 6-1, 9-2, 10-2, 10-3, 12-3, 17-2, 20-2, 22-1, 23-6, 33-1, 35-3, 38-1, 40-1, 50-1, 53-3, 54-2 and 59-1. These twenty problems represent a variety of partial differential operators and boundary conditions. Fifteen problems have Dirichlet boundary conditions, five of which are homogeneous. Fifteen of the domains are the unit square.

Each of the four numerical methods described above are applied to each of the subject population problems using the performance evaluation system of [Boisvert, et. al., 1979]. We use $n = 4, 8, 12, 20$ and 29 which involves from 100 to 3364 unknowns w_i . The computations are done on a VAX 11/780 computer with floating point accelerator using the UNIX FORTRAN compiler f77. Note that this experiment involves computing 400 solutions of elliptic problems.

4. Performance Analysis

We now consider the following hypothesis: *Scaling is essential for numerically solving the Hermite bicubic collocation equations.* To establish this hypothesis, we compare these methods pairwise using simple non-parametric analysis as follows:

Comparison	Interpretation
U/S vs U/U	Solve the unscaled equations with scaled partial pivoting (BAND GE) versus partial pivoting (LINPACK BAND)
S/U vs U/U	The scaled versus unscaled equations solved with partial pivoting
S/S vs S/U	Solve the scaled equations with scaled partial pivoting versus partial pivoting
S/S vs U/S	The scaled versus unscaled equations solved with scaled partial pivoting

The two methods of each pair are ranked on each problem using the maximum error at the grid points. For example, Figure 1 shows performance graphs of $\log(n+1)$ versus the

logarithm of the maximum error at the grid points for two problems. We see that method U/S is dramatically more accurate than method U/U; in fact, method U/U gives results which are so contaminated by round-off that they are totally unacceptable. By contrast, we see that methods S/S, S/U and U/S each give similar, accurate results; that is, the effect of either scaling the collocation equations or using scaled partial pivoting to solve them appears to remedy the problem present in method U/U. Note, however, in Figure 1 that for Problem 23-6 U/S and S/S give significantly better accuracy for one case. This means that our *a priori* scaling method was not nearly as good in this case as using scaled partial pivoting. These graphs typify the results obtained for the other problems.

We rank each pair of methods on each problem and compute average ranks for four different groups of problems: the ten with nonhomogeneous Dirichlet boundary conditions, the five with homogeneous Dirichlet boundary conditions, the five with mixed boundary conditions, and the entire subject population. An average rank of 1.00 means that the method is always the best whereas 2.00 means that it is always the worst. We obtain confidence levels on the observed differences using the Friedman, Kendall and Babington-Smith test [Hollander and Wolfe, 1973]. We summarize the results in Tables 1 - 4.

Table 1: Average rank of U/S vs U/U.

Group	Average Rank	Average Rank		Significance
		U/S	U/U	
Dir/Non (10)	1.00	2.00		99%
Dir/Hom (5)	1.70	1.30		< 80%
Mixed (5)	1.00	2.00		99%
Combined (20)	1.17	1.82		99%

Table 2: Average rank of S/U vs U/U.

Group	Average Rank	Average Rank		Significance
		U/S	U/U	
Dir/Non (10)	1.00	2.00		99%
Dir/Hom (5)	1.60	1.40		< 80%
Mixed (5)	1.00	2.00		97%
Combined (20)	1.15	1.85		99%

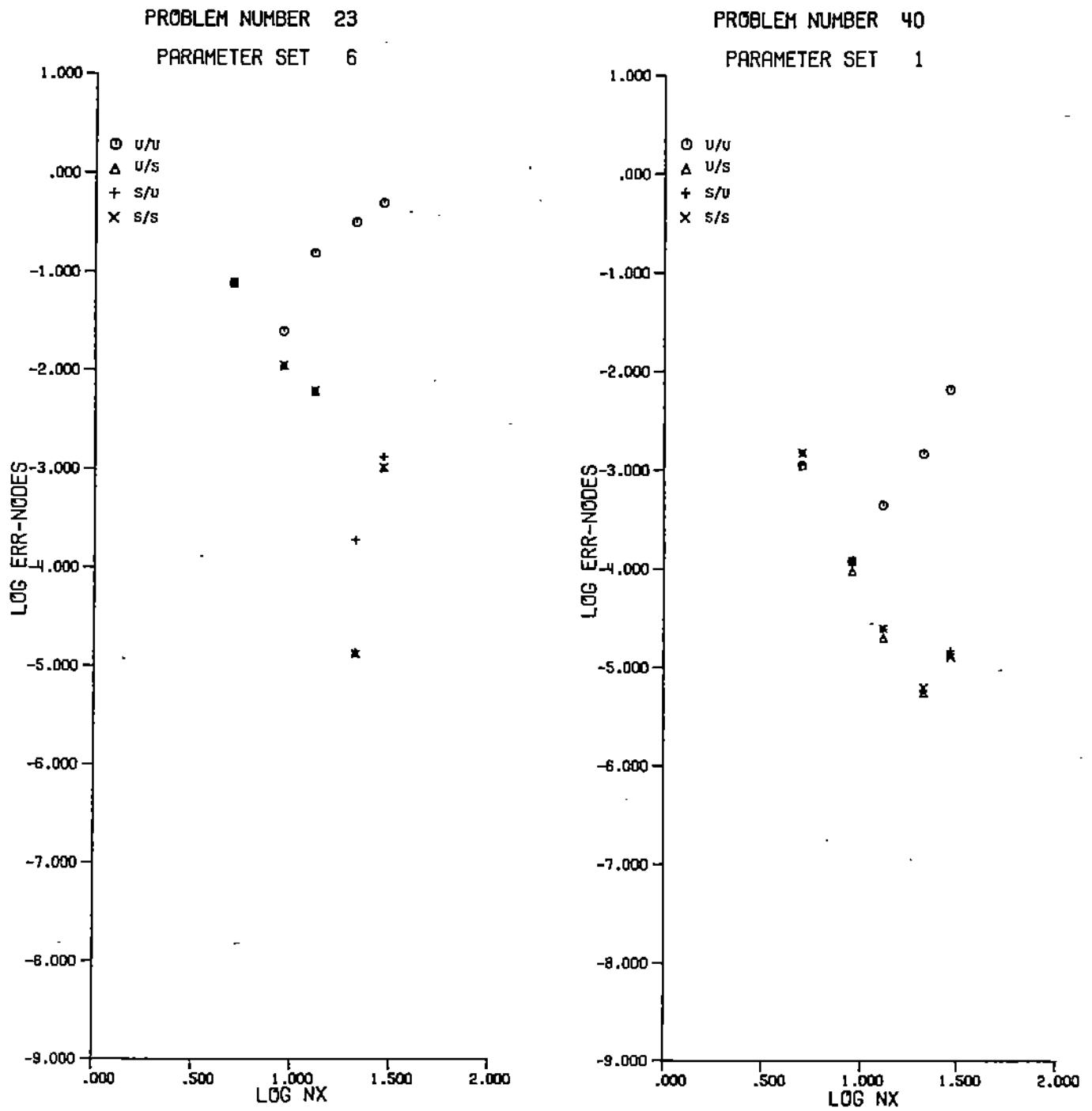


Figure 1: Graphs of the logarithm of $NX = n + 1$ versus the logarithm of the maximum error at the grid points for Problem 23-6 which has mixed boundary conditions and for Problem 40-1 which has nonhomogeneous Dirichlet boundary conditions.

Table 3: Average rank of S/S vs S/U.

Average Rank		U/S	U/U	Significance
Group				
Dir/Non	(10)	1.40	1.60	< 80%
Dir/Hom	(5)	1.70	1.30	< 80%
Mixed	(5)	1.50	1.50	< 80%
Combined	(20)	1.47	1.52	< 80%

Table 4: Average rank of S/S vs U/S.

Average Rank		U/S	U/U	Significance
Group				
Dir/Non	(10)	1.50	1.50	< 80%
Dir/Hom	(5)	1.60	1.40	< 80%
Mixed	(5)	1.20	1.80	80%
Combined	(20)	1.45	1.55	< 80%

For example, we see from Table 1 that comparing U/S versus U/U on the entire subject population gives a rank of 1.17 for U/S and a rank of 1.82 for U/U. The difference in rank is significant at the 99% level of confidence.

These experimental results strongly support our initial hypothesis, namely, that scaling is essential for numerically solving the Hermite bicubic collocation equations. We see from Figure 1 that the results obtained by scaling are significantly more accurate than those obtained by not scaling. Moreover, the data in Tables 1 - 4 demonstrate that the observed similarities or differences between methods are themselves statistically significant and not due merely to chance.

Our initial hypothesis can be stated more specifically in terms of the four methods considered here: method S/S is slightly more accurate than methods S/U and U/S which are all very much more accurate than method U/U. We believe that this data supports our hypothesis with a high level of statistical confidence.

Finally, we note from Tables 1 and 2 that problems with homogeneous Dirichlet boundary conditions are a significant special case. In this case, both HERMITE COLLOCATION and P3C1 COLLOCATION eliminate the boundary condition equations from the linear system during the discretization and before Gauss elimination. This suggests that the boundary

equations might be the key to understanding the severe round-off problems resulting from method U/U.

5. Scaling and the Boundary Equations

To further study the effects of round-off, we constructed a parameterized elliptic problem, Problem 59, whose solution is a bicubic for which Hermite bicubic collocation gives the exact solution except for round-off. Problem 59-1 is a Poisson problem with nonhomogeneous Dirichlet boundary conditions on the unit square.

Figures 2 and 3 show contour plots of the error for Problem 59-1 using S/U and U/U with $n = 8$. Figure 2 shows that if we solve the scaled equations using partial pivoting, then the error is rather randomly distributed and is of the order of machine precision, 10^{-6} . By contrast, we see from Figure 3 that if the unscaled equations are solved using merely partial pivoting, then the error in the interior of the domain is still on the order of 10^{-6} whereas the error on the boundary is on the order of 10^{-5} and is as large as 10^{-4} . Hence, essentially all of the round-off error occurs on the boundary; this is unexpected since the boundary conditions are Dirichlet and hence should be interpolated exactly. This is further evidence that the boundary equations are the key to understanding the round-off problems.

The relationship between the boundary equations and scaling is geometrical and can be seen by considering the order in which the equations are eliminated during Gauss elimination. Since the equations are associated with the collocation points, we can view the reordering of the equations produced by pivoting as a reordering of the collocation points themselves.

Figure 4 shows a typical example of the order of elimination resulting from solving Problem 59-1 using the unscaled collocation equations. We give the geometric ordering of the collocation equations before Gauss elimination and after Gauss elimination with scaled partial pivoting, partial pivoting and complete pivoting.

Figure 4 shows that the three pivoting strategies differ dramatically in the order in which they eliminate the interior and the boundary equations. Partial pivoting eliminates as

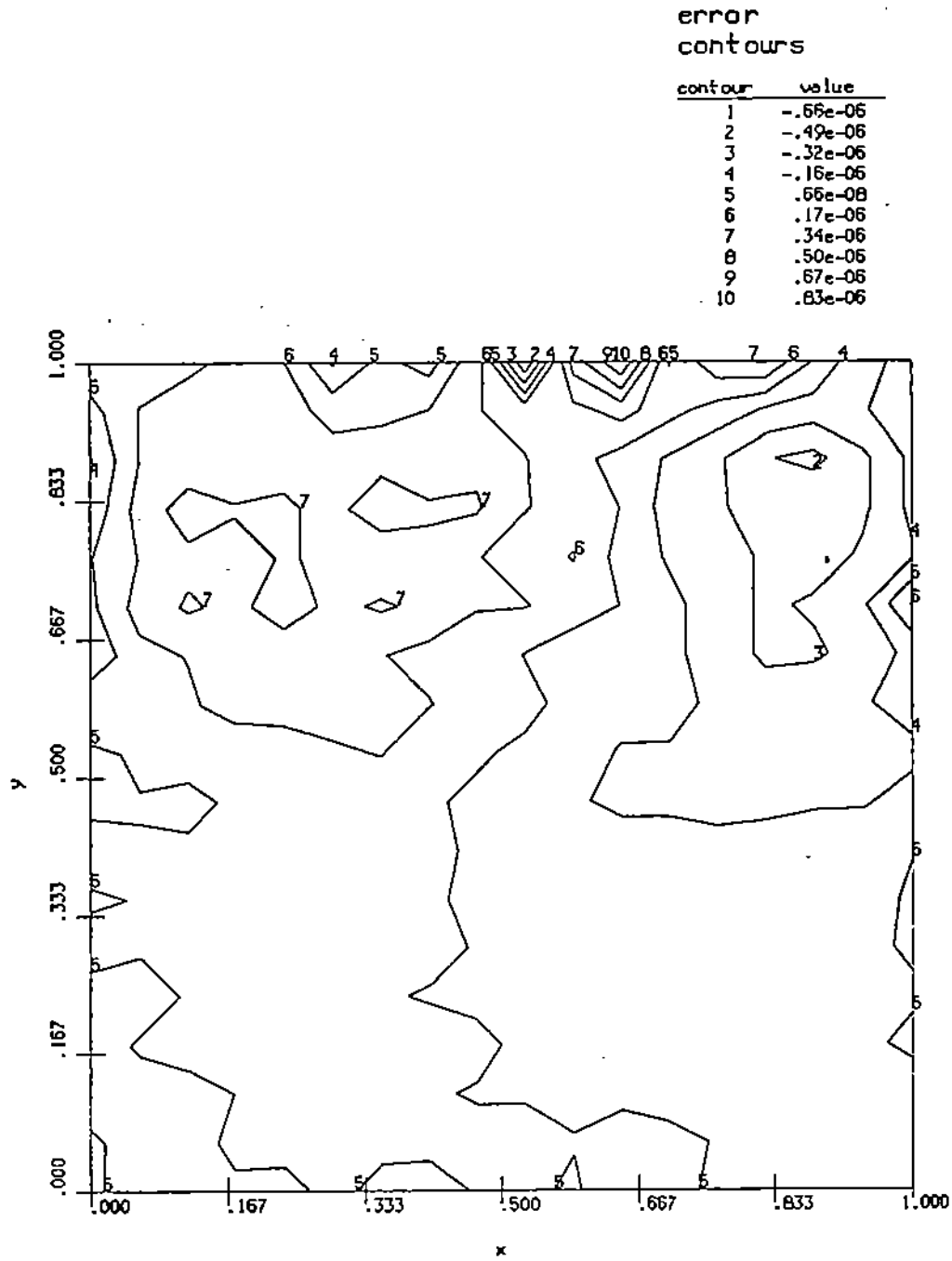


Figure 2: Contour plot of the error obtained by solving Problem 59-1 using S/U with $n = 8$. The error is due to round-off and is of the order of 10^{-6} which is machine precision.

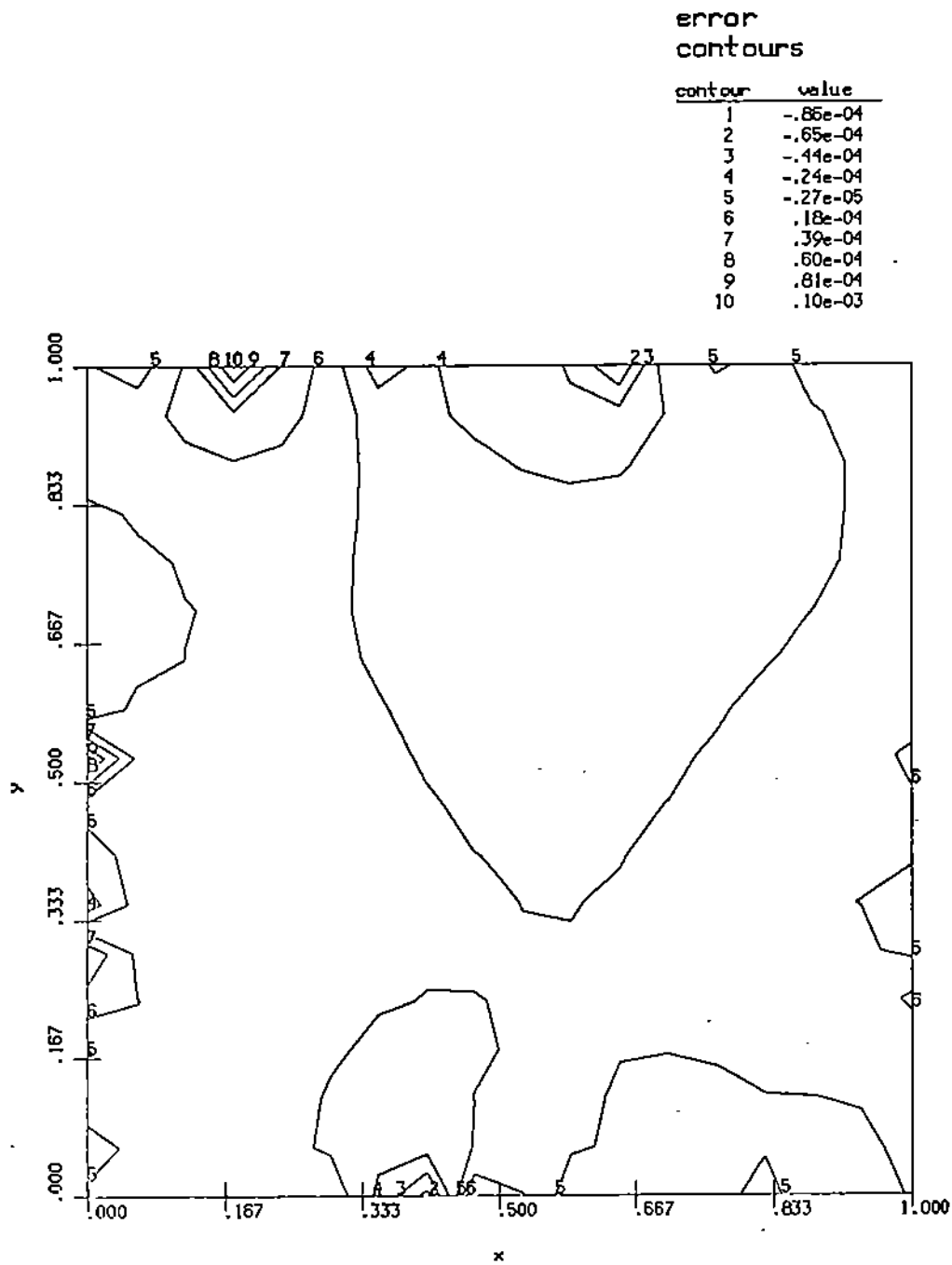


Figure 3: Contour plot of the error obtained by solving Problem 59-1 using U/U with $n=8$. The error in the interior of the domain is on the order of 10^{-6} whereas the error on the boundary is on the order of 10^{-5} and is as large as 10^{-4} .

many of the interior equations as possible before it must eliminate a boundary equation. Complete pivoting eliminates **all** of the interior equations before eliminating **any** of the boundary equations. Scaled partial pivoting weaves the elimination of the interior and boundary equations together; in fact, the reordering produced by scaled partial pivoting is essentially the Hermite Collorder ordering discussed in [Dyksen and Rice, 1983].

The above phenomena result from the inherent differences in magnitudes of the boundary and interior equations. In a typical elliptic Dirichlet problem, the coefficients of the boundary equations involve values of the basis functions and hence are $O(1)$. The interior equations, however, involve second derivatives of the basis functions and hence are $O(n^2)$. Thus, during simple partial pivoting and complete pivoting, the interior equations are chosen before the boundary equations as often as possible. As a result, the boundary condition information is not used until the last possible moment.

In practice, the lack of scaling using the original ordering results in the two dimensional analogue of numerically solving an ordinary differential equation from the inside out.

6. Scaling the Boundary Equations

There are two approaches to scaling the boundary equations. Since the scaling is required only for choosing the pivots, it need not be carried out explicitly, although to do so is a simple way to proceed. Thus, we can scale the boundary equations either **explicitly before** elimination or **implicitly during** elimination.

In order to determine an *a priori* scaling factor, we consider the Hermite bicubic U in the case in which all of the coefficients w_i are $O(1)$. If the domain is discretized with a uniform x and y spacing h_x and h_y , respectively, then a simple computation gives

$$(6.1) \quad U_{xx} + U_{yy} \sim (1 + h_y)(1 + 1/h_x)/h_x + (1 + h_x)(1 + 1/h_y)/h_y$$

and

Finite Element Ordering							
18	19	20	35	36	59	60	58
17	24	23	40	39	64	63	57
16	21	22	37	38	61	62	56
11	15	14	34	33	55	54	51
10	12	13	31	32	52	53	50
5	9	8	30	29	49	48	45
4	6	7	27	28	46	47	44
1	2	3	25	26	41	42	43

Scaled Partial Pivoting							
14	15	32	29	48	45	64	62
16	13	31	30	47	46	61	63
11	10	27	25	43	41	59	57
12	9	28	26	44	42	58	60
7	6	23	21	40	37	55	53
8	5	24	22	39	38	54	56
3	2	17	19	33	35	49	51
1	4	20	18	36	34	52	50

Partial Pivoting							
16	31	44	48	57	61	63	62
15	11	14	27	30	45	46	64
32	9	13	25	29	41	47	59
28	7	10	23	39	42	53	60
36	5	6	21	26	37	43	58
24	4	3	19	20	35	40	56
12	1	2	17	18	33	34	55
8	22	38	49	50	54	51	52

Complete Pivoting							
47	51	37	58	55	44	62	46
64	5	16	2	11	25	1	52
42	22	32	31	20	36	19	40
54	13	34	30	10	29	12	56
59	7	23	4	15	26	3	60
39	18	35	28	27	33	17	43
50	9	24	8	14	21	6	49
45	63	41	53	57	38	61	48

Figure 4: The geometric ordering of the collocation equations before Gauss elimination (upper left) and after Gauss elimination using scaled partial pivoting (upper right), partial pivoting (lower left) and complete pivoting (lower right).

$$(6.2) \quad \begin{aligned} U &\sim 1+h_y+h_x+h_x h_y \\ U_x &\sim 1/h_x+h_y/h_x+1+h_y \\ U_y &\sim 1/h_y+1+h_x/h_y+h_x \end{aligned}$$

Thus, for the model problem the interior equations look like (6.1) and the boundary equations look like (6.2).

With this in mind, we experimented extensively with many scaling factors, applying them to Problems 2-1, 22-1 and 59-1. We varied n and computed the maximum error at the grid points as well as the condition number of the coefficient matrix using the LINPACK routine SGBCO. For example, Table 5 summarizes the results for Problem 59-1 using the scale factor $1/h_x^2 + 1/h_y^2$. We see that scaling the boundary equations produces significant changes in both the error and the condition number.

Table 5: The effect of scaling on the condition number and the maximum error for Problem 59-1.

n	Number of Unknowns	Condition Number		Maximum Error	
		Scaled	Unscaled	Scaled	Unscaled
4	100	$1.4 \cdot 10^{+4}$	$5.1 \cdot 10^{+5}$	$9.5 \cdot 10^{-7}$	$1.1 \cdot 10^{-5}$
8	324	$5.3 \cdot 10^{+4}$	$6.9 \cdot 10^{+6}$	$9.5 \cdot 10^{-7}$	$8.0 \cdot 10^{-5}$
16	1156	$1.8 \cdot 10^{+5}$	$1.0 \cdot 10^{+8}$	$1.9 \cdot 10^{-6}$	$7.8 \cdot 10^{-4}$
22	2116	$3.2 \cdot 10^{+5}$	$3.5 \cdot 10^{+8}$	$3.7 \cdot 10^{-6}$	$1.5 \cdot 10^{-3}$
28	3364	$4.8 \cdot 10^{+5}$	$9.1 \cdot 10^{+8}$	$1.2 \cdot 10^{-5}$	$2.9 \cdot 10^{-3}$

Having experimented with these scaling factors, we propose the L_1 type scaling factor given in (6.1) [Skeel, 1979, 1980 and 1981]. It has a simple and natural analytical basis. Since from (6.2) $U = O(1)$, we multiply a boundary condition equation involving only u by (6.1) to make it the same size as the interior equations. Similarly, since from (6.2) $U_x = O(1/h_x)$ and $U_y = O(1/h_y)$, we scale a boundary equation involving u_x or u_y by the product of h_x or h_y , respectively, and (6.1). This proposed scaling method is in fact used by HERMITE COLLOCATION.

Although the above scaling method works well on a fairly large set of problems, we believe that it is not always practical to scale *a priori* the boundary equations to make them the same size as the interior equations. The scale factor in (6.1) is derived using a simple

model of the coefficients in the elliptic problem. The severe round-off phenomena observed above may occur again for problems in which either the coefficients in the partial differential operator are large or the coefficients in the boundary conditions are small at the collocation points. In such a case one would need to compute either the extreme values of these coefficient functions or perhaps the maximum L_1 norm of the interior equations to scale the boundary equations correctly. As a result, we conjecture that S/S is more reliable than S/U; that is, BAND GE is more reliable than its ancestor LINPACK BAND.

For example, Figure 5 shows a performance graph for Problem 22-1 which involves an operator with a large coefficient function. In this case, our *a priori* scaling method is clearly inferior to using scaled partial pivoting (with either the scaled or unscaled collocation equations).

As another example, consider the "scaled" Poisson problem

$$\begin{aligned} 10^k(u_{xx} + u_{yy}) &= f(x, y) \in \mathbf{R} = [0,1] \times [0,1] \\ u &= g(x, y) \in \partial\mathbf{R} \end{aligned}$$

where f and g are chosen so that $u = (x^3 + (xy)^2 + 2xy^3 + 1)/5$. If we vary k and solve this problem using S/S and S/U with $n = 21$ (1764 unknowns), we obtain the results given in Table 6. We again see that scaled partial pivoting is superior to our particular *a priori* scaling method.

Table 6: Maximum errors for the scaled Poisson problem.

k	Maximum Error	
	S/S	S/U
0	$3.1 \cdot 10^{-6}$	$2.6 \cdot 10^{-6}$
2	$1.4 \cdot 10^{-6}$	$1.2 \cdot 10^{-4}$
4	$1.5 \cdot 10^{-6}$	$1.0 \cdot 10^{-2}$
6	$6.3 \cdot 10^{-7}$	$8.5 \cdot 10^{-1}$
8	$1.7 \cdot 10^{-6}$	$2.0 \cdot 10^{+5}$

The data in Tables 1 - 4 do not provide any support for the conjecture that S/S is sometimes more reliable than S/U; the nature of those statistical tests masks this because the advantage of S/S shows up infrequently, only if the discretization error is close to round-off. In particular, there are 23 instances involving 11 problems in which the maximum relative

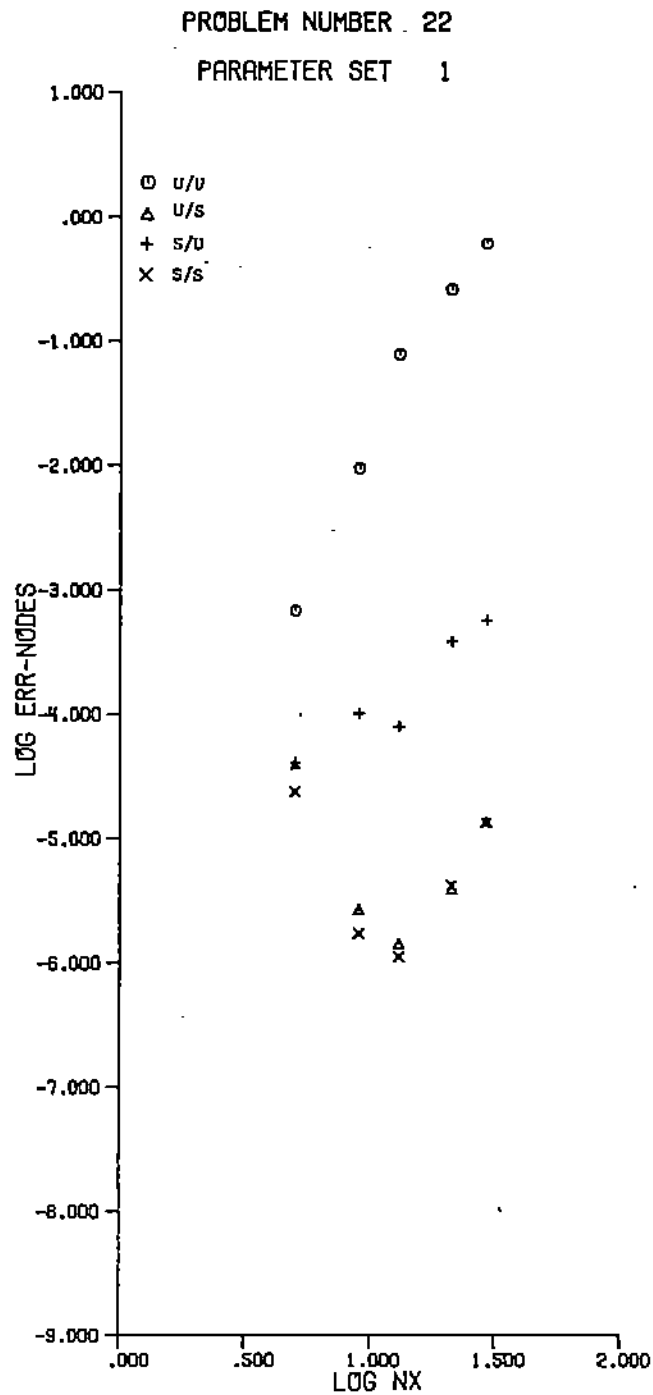


Figure 5: Graphs of the logarithm of $NX = n + 1$ versus the logarithm of the maximum error at the grid points for Problem 22-1 which involves an operator of the form $(wu_x)_x + (wu_y)_y = f$ where w is large.

error for either S/S or S/U is less than 10^{-5} . In twelve cases the errors differ by more than 10%, and S/S is more accurate in ten of these cases. For these twelve observations, the Sign test [Hollander and Wolfe, 1973] assures that S/S is more accurate than S/U with 98% confidence (p-value of 0.02).

The comparative work of different scaling methods is easily computed. For a grid of n^2 rectangles, there are $4n^2$ interior equations and $8n + 4$ boundary equations. The half bandwidth of the linear system is $4n + 7$ and there are at most 16 nonzero entries per interior equation and 8 nonzero entries per boundary equation. The comparative work for three scaling methods is given in Table 7. We see that each method requires much less work than Gauss elimination which is $O\left((4n + 7)^2 * 4(n + 1)^2\right) = O(n^4)$. Although scaling the boundary equations alone is the least amount of work, it is also probably the least reliable.

Table 7: Comparative work for three scaling methods.

Scaling Method	Work
Only boundary equations during discretization	$O(8 * (8n + 4)) = O(n)$
All equations during discretization	$O\left(16 * 4(n + 1)^2\right) = O(n^2)$
Scaled partial pivoting	$O\left((4n + 7) * 4(n + 1)^2\right) = O(n^3)$

The work estimates in Table 7 lead to an important observation. At each stage in scaled partial pivoting, the "scaled" entries below the diagonal are searched for a pivot. In the case of the collocation equations, this involves $4n + 7$ multiples even though there are at most 16 nonzero entries to examine. Clearly it is more efficient for the equations to be scaled during the discretization phase, before these relatively few nonzero entries are dispersed throughout the band. The resulting savings is an order of magnitude in the work of scaling. It might be that the overall best choice is to do the scaling of scaled partial pivoting during the discretization and then to use simple partial pivoting during the elimination.

In view of all of the above, we recommend using both the *a priori* scaling method described here along with scaled partial pivoting. Although the *a priori* scaling method is not

foolproof, it is simple to apply. Moreover, neither method of scaling requires any significant extra computation. For example, S/S takes on the average only 5% longer than S/U to solve the collocation equations with $n = 28$ (3364 unknowns).

Finally, we note that it is **not** the case that the Hermite bicubic collocation equations with unscaled boundary equations are inherently ill-scaled. In fact, we have observed that they can be solved accurately **without** scaling and **without** pivoting if one orders the equations and unknowns using the Hermite Collorder ordering given in [Dyksen and Rice, 1983].

7. Preservation of Geometric Integrity

The poor scaling of the collocation equations in their original form destroys the relationship between the geometry of the problem and the order of elimination. One hopes that the ellipticity of an elliptic problem should damp out errors, including round-off. However, destroying the geometry of the problem seems to ruin its ellipticity.

As a further example of this phenomenon, we consider the linear equations obtained from Problem 59-1 by using the standard 5-point star discretization modified to include the unscaled Dirichlet boundary equations. As in the case of Hermite bicubic collocation with the Hermite Collorder ordering, these 5-point star equations can be solved to machine precision **without** scaling and **without** pivoting. If the equations are solved with simple partial pivoting, then round-off dominates the computations as the grid is refined. We see that it is not only obviously inefficient to include the boundary equations haphazardly in the linear system, it is also dangerous. Note that this is done routinely in many finite element programs in structural engineering.

We also generated random row permutations and solved the equations using partial pivoting to see what effect if any this might have on the solution. In summary, we observed that the more the underlying geometry is perturbed, the larger the error becomes. This again suggests that the ordering used during Gauss elimination must maintain the geometric integrity of the elliptic problem. We believe that this is not particular to 5-point star or

Hermite bicubic collocation, and we conjecture that it is true for other numerical methods for elliptic problems.

References

- [1] R. F. Boisvert, E. N. Houstis and J. R. Rice, "A system for performance evaluation of partial differential equation software", *IEEE Trans. Software Engineering*, 5(1979), pp. 418-425.
- [2] S. D. Conte and C. W. de Boor, *Elementary Numerical Analysis*, McGraw Hill, New York, 1980.
- [3] J. J. Dongarra, C. B. Moler, J. R. Bunch and G. W. Stewart, *LINPACK Users' Guide*, SIAM, 1979.
- [4] W. R. Dyksen, "A new ordering scheme for the Hermite bicubic collocation equations", Purdue University, Computer Science Department Report CSD-TR 364, May 1, 1981.
- [5] W. R. Dyksen and J. R. Rice, "A new ordering scheme for the Hermite bicubic collocation equations", in *Elliptic Problem Solvers III*, (G. Birkhoff and A. Schoenstadt, eds.), Academic Press, 1984, pp. 467-480.
- [6] W. R. Dyksen, R. E. Lynch, J. R. Rice and E. N. Houstis, "The performance of the collocation and Galerkin methods with Hermite bi-cubics", *SIAM J. Numer. Anal.*, 21(1984), pp. 695-715.
- [7] M. Hollander and D. A. Wolfe, *Nonparametric Statistical Methods*, John Wiley, 1975.
- [8] E. N. Houstis, "Collocation methods for linear elliptic problems", *BIT*, 18(1978), pp. 301-310.
- [9] E. N. Houstis, R. E. Lynch, J. R. Rice and T. S. Papatheodorou, "Evaluation of numerical methods for elliptic partial differential equations", *J. Comput. Phys.*, 27(1978), pp. 323-350.
- [10] P. Percell and M. F. Wheeler, "A C^1 finite collocation method for elliptic problems", *SIAM J. Numer. Anal.*, 17(1980), pp. 605-622.
- [11] J. R. Rice, *Matrix Computations and Mathematical Software*, McGraw Hill, 1981.
- [12] J. R. Rice, E. N. Houstis and W. R. Dyksen, "A population of linear, second order, elliptic partial differential equations on rectangular domains, Parts 1 and 2", *Math. Comp.*, 36(1981), pp. 475-484.
- [13] J. R. Rice, *Numerical Methods: Software and Analysis*, McGraw Hill, 1983.
- [14] J. R. Rice and R. F. Boisvert, *Solving Elliptic Problems Using ELLPACK*, Springer-Verlag, 1984.

- [15] R. D. Skeel, "Scaling for numerical stability in Gaussian elimination", *J. ACM*, 26(1979), pp. 494-526.
- [16] R. D. Skeel, "Iterative refinement implies numerical stability for Gaussian elimination", *Math. Comp.* 35(1980), pp. 817-832.
- [17] R. D. Skeel, "Effect of equilibration on residual size for partial pivoting", *SIAM J. Numer. Anal.*, 18(1981), pp. 449-454.