

# A Gaussian Quadrature Method For The Numerical Solution of the Characteristic Initial Value Problem $u_{xy} = f(x, y, u)$

By J. T. Day

Gaussian quadrature methods have been used quite successfully in the numerical solution of ordinary differential equations (Cf. [1], [2]). We consider here a Gaussian quadrature method for the numerical solution of the characteristic initial value problem:

$$(1) \quad u_{xy} = f(x, y, u)$$

$$(1.1) \quad u(x, 0) = \sigma(x), \quad u(0, y) = \tau(y), \quad \sigma(0) = \tau(0)$$

$$(1.2) \quad 0 \leq x \leq a, \quad 0 \leq y \leq b$$

Throughout this discussion we shall assume sufficient regularity conditions for  $f$ ,  $\sigma$ , and  $\tau$  so as to insure that equations (4.1), (5), (6) have meaning in any context in which they are used. We also assume that values of  $u_x(0, y)$  and  $u_y(x, 0)$  as well as  $\sigma'(x)$  and  $\tau'(x)$  have been calculated: techniques on how this may be done are discussed in [5].

In order to apply quadrature methods for the numerical solution of the above initial value problem, we convert the differential equation (1) into a system of integral equations.

$$(2) \quad u(x_0 + h, y_0 + h) = u(x_0 + h, y_0) + u(x_0, y_0 + h) - u(x_0, y_0) + \int_{x_0}^{x_0+h} \int_{y_0}^{y_0+h} f(x, y, u(x, y)) dx dy$$

$$(3) \quad u_x(x_0 + h, y_0 + h) = u_x(x_0 + h, y_0) + \int_{y_0}^{y_0+h} f(x_0 + h, v, u(x_0 + h, v)) dv$$

$$(4) \quad u_y(x_0 + h, y_0 + h) = u_y(x_0, y_0 + h) + \int_{x_0}^{x_0+h} f(w, y_0 + h, u(w, y_0 + h)) dw$$

The numerical solution of (1) over a region  $D\{0 \leq x \leq a, 0 \leq y \leq b\}$  is to be carried out in a stepwise manner over a square mesh on  $D$ . The object of our method is that given  $u, u_x, u_y$  at  $(x_0, y_0), (x_0 + h, y_0), (x_0, y_0 + h)$  to calculate  $u, u_x, u_y$  at  $(x_0 + h, y_0 + h)$ .

For the evaluation of the double integral in (2) we shall use the cartesian product formula for the Gauss two-point rule on the interval  $[x_0, x_0 + h]$  (Cf. [3], [4]).

$$(4.1) \quad \int_{x_0}^{x_0+h} \int_{y_0}^{y_0+h} f(x, y, u(x, y)) dx dy = h^2/4 \cdot \sum_{k=1}^4 f(x_k, y_k, u(x_k, y_k)) + h^6/4320 \cdot \{D_x^4 f + D_y^4 f\}(x_0, y_0, u(x_0, y_0)) + \text{higher order terms}$$

(The symbols  $D_x$  and  $D_y$  mean  $D_x f = f_x + f_u u_x$ ;  $D_y f = f_y + f_u u_y$ .)

The points  $(x_k, y_k)$  are given by

$$(4.2) \quad \begin{aligned} x_1 &= x_0 + ph, & y_1 &= y_0 + ph \\ x_2 &= x_0 + qh, & y_2 &= y_0 + qh \\ x_3 &= x_0 + ph, & y_3 &= y_0 + qh \\ x_4 &= x_0 + qh, & y_4 &= y_0 + ph \end{aligned}$$

where  $p = (3 - \sqrt{3})/6, q = 1 - p$ . Hence, (2) becomes

$$(5) \quad \begin{aligned} u(x_0 + h, y_0 + h) &= u(x_0 + h, y_0) + u(x_0, y_0 + h) - u(x_0, y_0) \\ &+ h^2/4 \sum_{k=1}^4 f[x_k, y_k, u(x_k, y_k)] + O(h^6). \end{aligned}$$

Assuming that  $u_x$  and  $u_y$  have been calculated in the previous step, either from the given initial data or by a method such as indicated below, we have by means of Taylor expansions, the following estimate for  $u$  at the points  $(x_k, y_k)$ .

$$(6) \quad \begin{aligned} u(x_0 + sh, y_0 + th) &= u(x_0, y_0)(1 - s^2 - t^2) \\ &+ s^2u(x_0 + h, y_0) + t^2u(x_0, y_0 + h) \\ &+ hs(1 - s)u_x(x_0, y_0) + ht(1 - t)u_y(x_0, y_0) \\ &+ sth^2f[x_0, y_0, u(x_0, y_0)] + O(h^3). \end{aligned}$$

Substituting the values of  $u(x_k, y_k)$  calculated by (6) into (5), we obtain an approximate value for  $u(x_0 + h, y_0 + h)$  with an error of order  $h^5$ .

A bound for the local truncation error, assuming that the results from the previous steps are exact, is given by

$$|u - u^c| \leq h^5L[M/12 + L(1 + N)/6] + h^6A/2160 + O(h^7).$$

Here  $u^c$  denotes the approximate value of  $u(x_0 + h, y_0 + h)$ .

$$L = \text{Sup}(|f_x|, |f_y|, |f_u|)$$

$$M = \text{Sup}(|u_{xxx}|, |u_{yyy}|)$$

$$A = \text{Sup}(|D_x^4f(x, y, u)|, |D_y^4f(x, y, u)|)$$

$$N = \text{Sup}(|u_x|, |u_y|).$$

The suprema are taken over  $x, y$  where  $(x, y) \in R$ .

By the use of techniques discussed in [8] and [9], it can be shown that the cumulative error for the method under discussion has the following bound at a point  $(x, y)$  of the square mesh on  $D$ :

$$|e(x, y)| \leq (I_0(2\sqrt{Lxy}) - 1) \cdot [M/12 + L(1 + N)/6] \cdot h^3.$$

Here  $I_0$  is the modified Bessel function of order zero (i.e.,  $I_0(x) = J_0(ix)$ ).

$L, M, N$  are defined as above except that the suprema are taken over points  $(\xi, \eta)$  of the rectangle  $0 \leq \xi \leq x, 0 \leq \eta \leq y$ .

This bound takes into account only the accumulative effect of the order  $h^5$

errors, with the additional assumption that  $u_x$  and  $u_y$  have been calculated correctly to order  $h^3$  at every point of the mesh.

Since we now have an approximate value for  $u$  at  $(x_0 + h, y_0 + h)$ , we can calculate  $u_x$  and  $u_y$  at  $(x_0 + h, y_0 + h)$  by means of the trapezoidal rule or by other equally spaced numerical integration schemes applied to (3) and (4). In the case of the trapezoidal rule, we obtain an error of order  $h^3$  for  $u_x$  and  $u_y$ ; moreover, we note that the three values of  $f$  used in calculating  $u_x$  and  $u_y$  may be used in succeeding calculations.

We note that the number of evaluations of  $f$  over each subsquare is eight, three of which may be used in succeeding calculations, if storage limitations permit. The number of evaluations used in [5] for the solution of the more general problem  $u_{xy} = f(x, y, u, u_x, u_y)$  to the same order of accuracy is 15.

We believe that this method would also be of value in obtaining starting values for methods such as those discussed in [6]. With this point of view, we have performed calculations for the following example.

The equation of Liouville [7]:

$$u_{xy} = e^{2u}$$

with initial conditions

$$u(x, 0) = x/2 - \log(1 + e^x)$$

TABLE 1  
*Errors\**

$x \backslash y$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
0.25	0.29	1.10	1.74	1.97	1.91	1.68	1.46	1.28
0.50	1.10	3.50	5.54	6.38	6.20	5.58	4.92	4.31
0.75	2.23	6.93	11.28	13.36	13.30	12.17	10.78	9.48
1.00	3.50	11.01	18.80	23.25	23.85	22.21	19.82	17.49
1.25	4.66	15.16	27.46	35.89	38.32	36.53	32.97	29.17
1.50	5.54	18.80	36.34	50.69	56.85	55.91	51.24	45.62
1.75	6.11	21.53	44.35	66.39	79.10	80.95	75.87	68.16
2.00	6.36	23.23	50.67	81.51	103.96	111.89	108.08	98.50
2.25	6.36	23.93	54.84	94.41	129.59	148.21	148.99	138.71
2.50	6.18	23.84	56.85	103.95	153.62	188.29	199.08	190.95
2.75	5.90	23.18	57.05	109.72	173.74	229.35	257.72	247.19
3.00	5.57	22.19	55.87	111.87	188.26	267.65	322.41	338.42
3.25	5.24	21.04	53.81	111.05	196.57	299.57	388.71	433.70
3.50	4.91	19.81	51.21	108.02	199.02	322.36	450.61	539.04
3.75	4.60	18.63	48.43	103.64	196.73	335.15	502.23	647.24
4.00	4.30	17.47	45.58	98.43	190.87	338.33	538.98	748.34
4.25	3.99	16.32	42.72	92.86	182.70	333.66	559.13	832.45
4.50	3.70	15.27	40.04	87.31	173.33	323.36	563.79	892.34
4.75	3.45	14.33	37.58	81.99	163.61	309.51	555.83	925.32
5.00	3.21	13.42	35.22	76.86	153.80	293.61	538.50	932.72
5.25	2.97	12.55	33.01	72.04	144.30	276.93	515.21	919.17
5.50	2.78	11.80	31.00	67.63	135.38	260.44	488.79	890.49
5.75	2.60	11.12	29.19	63.57	127.07	244.53	461.24	852.17
6.00	2.43	10.45	27.49	59.80	119.36	229.49	433.84	808.59

\* All errors in the table are to be multiplied by  $10^{-8}$ .

$$u(0, y) = y/2 - \log(1 + e^y).$$

The solution of this problem is  $u(x, y) = (x + y)/2 - \log(e^x + e^y)$ .

Taking  $h$  as 0.05, we have written a program in Fortran for the CDC 1604 computer at the University of California, San Diego and have found the following errors for the computation of  $u$  at the points  $(x, y)$  given below. By error we mean here the relative error, i.e.,

$$\text{error} = |(\text{true value} - \text{approximate value})/\text{true value}|.$$

As a concluding remark, we wish to point out that the techniques used here can be applied to the more general equation  $u_{xy} = f(x, y, u, u_x, u_y)$ . In this case, one could use the quadrature formula used above and the Moore-Runge-Kutta method to estimate the values of  $u, u_x, u_y$  at the quadrature evaluation points.

I am especially indebted to Prof. P. C. Hammer and A. H. Stroud for many discussions on the general techniques of numerical integration in more than one variable.

University of California, San Diego  
La Jolla, California

1. P. C. HAMMER & J. W. HOLLINGSWORTH, "Trapezoidal methods of approximating solutions of differential equations," *MTAC*, v. 9, 1955, p. 92-96.
2. D. MORRISON & L. STOLLER, "A method for the numerical integration of ordinary differential equations," *MTAC*, v. 12, 1958, p. 269-272.
3. P. C. HAMMER & A. W. WYMORE, "Numerical evaluation of multiple integrals I," *MTAC*, v. 11, 1957, p. 59-67.
4. J. ALBRECHT & L. COLLATZ, "Zur numerischen Auswertung mehrdimensionaler Integrale," *Z. Angew. Math. Mech.*, v. 38, 1958, p. 8.
5. R. H. MOORE, "A Runge-Kutta procedure for the Goursat problem in hyperbolic partial differential equations," *Arch. Rational Mech. Anal.*, v. 7, 1961, p. 37-63.
6. W. TÖRNIG, "Zur numerischen Behandlung von Anfangswertproblemen partieller hyperbolischer Differentialgleichungen zweiter Ordnung in zwei unabhängigen Veränderlichen," *Arch. Rational Mech. Anal.*, v. 4, 1960, p. 428-445.
7. C. JORDAN, *Cours d'Analyse*, Gauthier-Villars, Paris, Third Edition, v. 3, 1915, p. 369-371.
8. W. WALTER, "Fehlerabschätzungen bei hyperbolischen Differentialgleichungen," *Arch. Rational Mech. Anal.*, v. 7, 1961, p. 249-272.
9. P. HENRICI, *Discrete Variable Methods in Ordinary Differential Equations*, John Wiley and Sons, New York, 1962.

## Gauss Elimination for Singular Matrices

By George Shapiro

**1. Introduction.** Several variations of a method of successive elimination, associated with the name of Gauss, are frequently used to solve systems of linear equations or to invert a matrix. Multiplication of the inverse matrix by the determinant of the original matrix (which is readily available as an intermediate result of the Gauss elimination) yields the adjoint of the original matrix.

Recently, the use of modular or residue class arithmetic systems for high-speed computers has been considered [1]. In such systems, an integer is represented

---

Received November 13, 1962, revised March 18, 1963. The research leading to this paper was sponsored by the Electronic Technology Laboratory of the USAF Aeronautical Systems Division.