Compressible Fluid Flow and the Approximation of Iterated Integrals of a Singular Function*

By P. L. Richman

Abstract. A computer implementation of Bergman's solution to the initial value problem for the partial differential equation of compressible fluid flow is described. This work necessitated the discovery of an efficient approximation to the iterated indefinite integrals of an implicitly defined real function of a real variable with a singularity which is not included in the possible domains of integration. The method of approximation used here and the subsequently derived error bounds appear to have rather general applications for the approximation of the iterated integrals of a singular function of one real variable.

Acknowledgment. I thank Professors Bergman and Herriot for their valuable help and guidance in this work, and for their permission and encouragement for me to publish these results.

1. Introduction. Of interest here are the initial and boundary value problems for the partial differential equation describing the two-dimensional, irrotational, steady, free from turbulence, adiabatic flow of an ideal, inviscid, compressible fluid. The first task in devising a numerical procedure for solving such problems is that of finding a *constructive* mathematical solution to the problem. For certain subsonic domains in the physical plane, a constructive solution to the boundary value problem can be found in [B.2], [B.3], and [B.4]. It is given there as an infinite series of orthogonal polynomials which converges only in (a part of) the subsonic region. In order to continue this solution into the supersonic region, Bergman suggests using this (explicit) subsonic solution to set up an initial value problem of mixed type. The solution to this initial value problem, as given in [B.2], may then be valid in some part of the supersonic region. (The particular solutions to the flow equation which are used here and in [B.2] were obtained independently by Bergman and Bers-Gelbart.) Whether this continuation will lead to a closed, meaningful flow is an open question.

Even after such constructive solutions are found, there is much to be done before actual computation can be carried out. In this paper, we deal with the solution of the initial value problem of mixed type. It is in this connection that the iterated integrals of a singular function arise (the singularity being near to, but not in, the possible domains of integration).

These procedures for generating flow patterns are different from that using Bergman's integral operator [B.1] and the examples of this paper are different from those obtained by Stark [S], using this integral operator. See also Ludford [L] and Finn and Gilbarg [F-G].

Received July 15, 1968.

^{*} This work was supported in part by N.S.F. GP 5962, O.N.R. 225(37), and Air Force AF1047-66, and used as part of the author's thesis.

In Section 2, the initial value problem and its solution are presented.

In Section 3, we approximate the iterated integrals arising in the solution of Section 2 for the special case in which the fluid under consideration is air.

In Section 4, the methods of approximation used in Section 3 are generalized to cover an arbitrary fluid.

In Section 5, a numerical procedure for generating the solution to the initial value problem is given briefly, and a sample flow pattern is included. A new relation between the speed, v(H), and the iterated integrals, $s_m(H, H_0)$, is also given.

In Section 6, a priori absolute error bounds are derived for the truncation and function approximation errors. To illustrate the effectiveness of these bounds, we analyze the error involved in computing, by our method, the well-known Ringleb solution [R].

Our principal results for numerical analysis are the development of an efficient method for approximating the iterated indefinite integrals of a singular function (Sections 3, 4) and the derivation of a tight error-bound for the error arising in such an approximation, excluding roundoff (Section 6). In comparing our method of approximation with a straightforward polynomial approximation technique, we find that our method offers

(1) considerably more accuracy for the number of arbitrary coefficients used in the approximation to the function (l(H)) to be iteratively integrated (see (2.6) and/or (3.1) for a definition of the iterated integrals), and

(2) better numerical properties; our method avoids a fit to l(H) with large coefficients of alternating signs so that we can use single precision for our computations, and our method involves considerably smaller powers of a certain variable (see Table 3.2) so that we avoid overflow/underflow problems.

These advantages are obtained by making effective use of available information about the singularity of l(H).

2. The Initial Value Problem. The partial differential equation describing the flow of an inviscid, ideal, compressible fluid is nonlinear when considered in the physical plane (x, y-plane). However, when transformed into the so-called hodo-graph plane $(H, \theta\text{-plane})$, this equation becomes a linear one, namely (see [B-H-K] for a description of the physical problem and explanation of the hodograph transformation):

(2.1)
$$\partial^2 \Psi / \partial H^2 + l(H) \partial^2 \Psi / \partial \theta^2 = 0, \quad l(H) = (1 - M^2) / \rho^2,$$

where

(2.2)
$$H \equiv H(v) = \int_{v_1}^{v} \frac{\rho(s)}{s} \, ds \, ,$$

(2.3)
$$\rho \equiv \rho(v) = \left\{1 - \frac{1}{2}(k-1)(v/a_0)^2\right\}^{1/(k-1)},$$

(2.4) $M = v / \{ a_0^2 - \frac{1}{2}(k-1)v^2 \}^{1/2}.$

Here θ is the angle which the velocity vector forms with the positive direction of the x-axis, v is the speed, $\Psi(H, \theta)$ is the stream function, M is the Mach number, ρ is the density, v_1 is the speed when M = 1 (i.e., the speed on the sonic line), k is a constant depending on the fluid, and a_0 is a conveniently chosen constant.

We shall describe a numerical procedure for solving the initial value problem in which the stream function, $\Psi(H_0, \theta) = f(\theta)$, and its derivative,

$$\left. \frac{\partial \Psi(H,\theta)}{\partial H} \right|_{H=H_0} = g^{(1)}(\theta) ,$$

are specified on an arbitrary line, $H = H_0$. The basis for this procedure is provided by the following

THEOREM 2.1 (SEE [B.2, p. 895]). Let α and β satisfy $\alpha < \beta < H(a_0(2/(k-1))^{1/2})$. Suppose that for $|\theta| \leq \theta_1$ and a given $H_0 \in [\alpha, \beta]$ we have

(2.5)

$$\Psi(H_0, \theta) = \sum_{n=0}^{\infty} C_n \theta^n \equiv f(\theta) ,$$

$$\left. \frac{\partial \Psi(H, \theta)}{\partial H} \right|_{H=H_0} = \sum_{n=0}^{\infty} n D_n \theta^{n-1} \equiv g^{(1)}(\theta) ,$$

where the series $\sum C_n \theta^n$ and $\sum D_n \theta^n$ converge uniformly and absolutely for $|\theta| \leq \theta_1$. Suppose that $|l(H)| \leq c^2, 0 < c < \infty$, for $H \in [\alpha, \beta]$. Let us define functions $s_m(H, H_0)$ by $s_0(H, H_0) = 1$, $s_1(H, H_0) = H - H_0$, and for $m = 2, 3, \cdots$

(2.6)
$$s_m(H, H_0) = \int_{H_0}^H \int_{H_0}^{H_1} l(H_2) s_{m-2}(H_2, H_0) dH_2 dH_1$$

Then, for H and θ satisfying $|\theta| + c|H - H_0| \leq \theta_1$ and $H \in [\alpha, \beta]$,

(2.7)
$$\Psi(H,\theta) = \sum_{j=0}^{\infty} (-1)^{j} \{ s_{2j}(H,H_0) f^{(2j)}(\theta) + s_{2j+1}(H,H_0) g^{(2j+1)}(\theta) \}$$

is the unique (analytic) solution of (2.1) satisfying (2.5). Here $f^{(i)} \equiv d^{i}f/d\theta^{i}$ and $g^{(i+1)} \equiv d^{i}g^{(1)}/d\theta^{i}$.

It is easy to check that (2.7) satisfies (2.1) and (2.5). For a proof of (absolute and uniform) convergence of (2.7) see [B.2, p. 896]. (However, there is an incorrect specification of the domain of convergence in this reference. The domain stated there is $\{(H, \theta): |\theta| + c|H - H_0| \leq \theta_1\}$, whereas the domain of convergence actually established by his proof is

$$\{(H,\theta)\colon |\theta|+c|H-H_0|\leq \theta_1 \text{ and } |H-H_0|\leq H_1\}.$$

The constraint $|H - H_0| \leq H_1$ corresponds to our constraint, $H \in [\alpha, \beta]$.)

The domain of convergence guaranteed by this theorem is diamond shaped (possibly truncated). If the initial conditions are specified by a Fourier series instead of a power series, then a theorem similar to this one can be proved. In that case, the domain of guaranteed convergence would be rectangular.

In any numerical evaluation of the right-hand side of (2.7) we have to approximate all functions in a convenient way, and we must truncate the series. We shall denote approximation functions by enclosing the function name in brackets. In this manner (2.7) becomes

(2.8)
$$[\Psi_n](H, H_0, \theta) = \sum_{j=0}^n (-1)^j \{ [s_{2j}](H, H_0) [f^{(2j)}](\theta) + [s_{2j+1}](H, H_0) [g^{(2j+1)}](\theta) \}.$$

P. L. RICHMAN

The approximation, $[\Psi_n]$, to Ψ depends on H_0 , whereas Ψ does not. The following remarks about $f^{(2j)}$ will apply to $g^{(2j+1)}$ as well. In general, obtaining approximations $[f^{(2j)}]$, for $j = 0, 1, \dots, n$, is not difficult. In fact, in the usual application of this procedure $f^{(2j)}$ will be defined in terms of functions customarily available on computers, such as sine, cosine, etc., and it will be possible to calculate $f^{(2j)}$ to almost full machine accuracy. In such cases the fact that we are really calculating a $[f^{(2j)}]$ is somewhat obscured by our ability to express it, in current programming languages, in precisely the form of its formal definition. For example, the Algol statement to calculate an approximation to $h(x) = \sin x$ is just " $h: = \sin (x)$." However, when only [f], and not f, is known (perhaps as the result of solving the boundary value problem alluded to earlier in this paper) a severe error is incurred. This is why we keep track of $f^{(2j)} - [f^{(2j)}]$ in the analysis of Section 6.

The values of $[f^{(2j)}](\theta)$ may be derived from an approximation, [f]. For example, if f is given as in (2.5), we can truncate that series to obtain [f]. We can then use an iterative synthetic division scheme to evaluate $[f]^{(2j)}(\theta)/(2j)!$, for $j = 0, 1, \dots, n$. (Note that $[f]^{(m)}$ denotes the *m*th derivative of [f] and $[f^{(m)}]$ denotes an approximation to $f^{(m)}$; $[f]^{(m)}$ need not be a very good $[f^{(m)}]$.) Of course the error of $[f]^{(2j)}$ incurred by such a procedure increase as j increases. However, if (any suitable norm of) the $f^{(2j)}$, considered as a function of j, does not increase too rapidly for $j \leq n$, then the absolute error of $s_{2j}f^{(2j)}$ will not increase as j grows and remains $\leq n$. This is because $s_m(H, H_0) \to 0$ rapidly as $m \to \infty$, since, as indicated in [B.2],

(2.9)
$$|s_m(H, H_0)| \leq \frac{{\delta_m}^{-1}}{m!} c^m |H - H_0|^m$$
,

where $\delta_m = c$ for m odd and $\delta_m = 1$ for m even, and c is the constant in Theorem 2.1. (See Section 6 for a further discussion of this point.)

The determination of $[s_m]$ presents more challenging problems. Because of the nature of l, an exact formula for s_m has not been found. The numerical procedure which evaluates $[\Psi_n]$ will be used to trace the streamlines, $\Psi(H, \theta) = \text{const.}$ Such curves, when transformed into the x, y-plane, describe the fluid flow. This means that many evaluations of $[\Psi_n]$ will be required, and so the [l] for the l in (2.6) must be chosen to yield an efficient scheme. In the next section we derive such an approximation to l and thus to s_m for the special case in which the fluid under consideration is air. In this case

(2.10)
$$k = 1.4, v_1 = (5/6)^{1/2},$$

and we choose $a_0 = 1$ (see (2.3) and (2.4)). The general case, where k > 1, is dealt with in Section 4.

3. The Integrals $s_m(H, H_0)$ and Their Approximation for k = 1.4. We will consider H_0 satisfying $H_0 < .25125 \cdots = H(\sqrt{5}) \equiv p$, since as $H \to p$ the Mach number, M(H), approaches infinity. A major problem in this implementation was the construction of an approximation, [l], to l over some subinterval of $(-\infty, p)$ which would allow a relatively simple expression for $[s_m]$. The approximation of [B-H-K] was not satisfactory for our purposes. It consisted of two tenth-degree polynomial approximations, one for the region [-1, 0] and the other for [0, .2]. The maximum absolute error of this combined approximation was 2.6×10^{-3} . In

[B-H-K], H_0 was fixed at zero, and so their approximation led to two expressions for $[s_m](H, 0)$, one valid in [-1, 0], and the other in [0, .2]. In our work H_0 is arbitrary, so we must have a single representation for [l] and $[s_m](H, H_0)$. As indicated in [B-H-R, p. 8], more precise knowledge of the singularity of l facilitates the determination of such a single representation. We observe that, for k = 1.4, l has the expansion,

(3.1)
$$l(H) = \sum_{j=0}^{\infty} b_j (p - H)^{2(j-6)/7}.$$

Thus *l* has a singularity at $p = .25125 \cdots$ of order 12/7. (The first 43 coefficients, b_0, \cdots, b_{42} , are given in [B-H-R, p. 9].) Equation (3.1) follows from (2.1) to (2.4) by substituting 5(1 - t) for s^2 so that

(3.2)
$$H = \int_{(5/6)^{1/2}}^{v} (1 - .2s^2)^{2.5} \frac{ds}{s} = -\frac{1}{2} \int_{5/6}^{\tau} \frac{t^{2.5} dt}{1 - t} = -\frac{1}{2} \int_{5/6}^{\tau} (t^{5/2} + t^{7/2} + \cdots) dt,$$

(3.3)
$$p - H = \frac{\tau^{7/2}}{\epsilon} + \frac{\tau^{9/2}}{9} + \cdots$$

(3.4)
$$l(H) = \frac{6\tau - 5}{\tau^6}$$
.

An approximation of the form

(3.5)
$$[l](H) = \sum_{j=0}^{7} a_j (p - H)^{(2j-12)/7}$$

was found for $H \in [-1, .22]$ by using the Remez algorithm, as adapted for the B5500 computer by Golub and Smith [G-S], to calculate the best values, in the Chebyshev sense, for a_0, a_1, \dots, a_7 . These values are listed in Table 3.1. The Remez algorithm reported the maximum error of this approximation to be

(3.6)
$$\max_{-1 \le H \le .22} |l(H) - [l](H)| = 4.10533 \times 10^{-5}.$$

TABLE			.1.	
m				F 73

The	coefficients	for	[l]	•
			L. 1	-

\overline{j}	aj		
0	-0.1505866818		
1	-0.4018655347		
2	2.0945191543		
3	-5.8821787341		
4	10.9583158033		
5	-10.7524447788		
6	5.9416272229		
7	-0.8198101027		

P. L. RICHMAN

It should be noted that this maximum error is not a mathematical bound; it is a computed value, subject to errors that are not analyzed by Golub and Smith. (However, our experience with this [l](H) has given us much faith in (3.6).)

This approximation is to be contrasted with that of [B-H-K], where 22 arbitrary coefficients (instead of 8) are used, the interval [-1, .2] (instead of [-1, .22]) is used, a maximum error of 10^{-3} (instead of 10^{-5}) is attained, two separate representations of $[s_m]$ (instead of one) are required, and the coefficients chosen by Remez for [l] are large and have alternating signs, requiring the evaluation of their $[s_m]$'s to be done in double precision. There is one more advantage to our approximation, which we give after the following representation theorem for $[s_m]$.

THEOREM 3.1. Let the $[s_m]$ be connected by the recurrence relation

(3.7)
$$[s_m](H, H_0) = \int_{H_0}^H \int_{H_0}^{H_1} [l](H_2)[s_{m-2}](H_2, H_0) dH_2 dH_1 \quad \text{for } m \ge 2 ,$$

where

$$(3.8) [s_0](H, H_0) = 1, [s_1](H, H_0) = H - H_0,$$

(3.9)
$$[l](H) = \sum_{j=0}^{N-1} a_j (p-H)^{(2j-12)/7}, \text{ and } 7 \leq N < \infty.$$

Then $[s_m]$ can be expressed as

(3.10)
$$[s_m](H, H_0) = \sum_{j=0}^{mN} c_{m,j} (p-H)^{j/7} \text{ for } m = 0, 1, 2, \cdots,$$

where $c_{m,1} = c_{m,3} = c_{m,5} = 0$ for all m and $c_{0,0} = c_{1,0} = -c_{1,7} = 1$ and $c_{1,j} = 0$ for $j \neq 0, 7$. The $c_{m,j}$ and $c_{m-2,j}$ are connected by the following recurrence relations:

(3.11)
$$c_{m,j} = -7\beta_{m,j-2}/j \text{ for } j = 1, 2, \cdots, mN \text{ with } j \neq 7,$$

(3.12)
$$c_{m,7} = \sum_{j=0}^{m/r-2} \beta_{m,j} (p - H_0)^{(j-5)/7}$$

(3.13)
$$c_{m,0} = -\sum_{j=2}^{mN} c_{m,j} (p - H_0)^{j/7},$$

where

(3.14)
$$\beta_{m,1} = \beta_{m,3} = \beta_{m,5} = 0$$
,

and for $j = 0, 2, 4, 6, 7, 8, \dots, mN - 2$, with entire(x) denoting the greatest integer in the real number x,

(3.15)
$$\beta_{m,j} = \frac{7}{5-j} \sum_{n=n_1(j)}^{n_2(j)} a_n c_{m-2,j-2n},$$

where

$$n_1(j) \equiv \text{entier}(\frac{1}{2} + \frac{1}{2} \max \{0, j - (m - 2)N\})$$

and

$$n_2(j) \equiv \operatorname{entier}(\frac{1}{2} \min \{j, 2N - 2\})$$

Proof. Equation (3.10) holds for m = 0, 1. We proceed by induction, assuming that (3.10) to (3.15) hold for m - 2 and proving them for m. We have

(3.16)
$$[s_{m-2}](H, H_0)[l](H) = \sum_{j=0}^{N-1} a_j (p-H)^{(2j-12)/7} \sum_{n=0}^{(m-2)N} c_{m-2,n} (p-H)^{n/7} \\ = \sum_{j=0}^{mN-2} \alpha_{m,j} (p-H)^{(j-12)/7} ,$$

where, for $j = 0, 1, \dots, mN - 2$,

(3.17)
$$\alpha_{m,j} = \sum_{n=n_1(j)}^{n_2(j)} a_n c_{m-2,j-2n} \, .$$

Since (3.16) is to be integrated, we must show that $\alpha_{m,5} = 0$, so that the term $\alpha_{m,5}(p-H)^{-1}$ drops out of (3.16) and no log (p-H) terms enter. Part of our induction hypothesis is $c_{m-2,5-2n} = 0$ for n = 0, 1, 2, and so

(3.18)
$$\alpha_{m,5} = \sum_{n=n_1(5)}^{n_2(5)} a_n c_{m-2,5-2n} = 0.$$

The rest of the proof follows as a formal calculation.

This procedure for approximating a singular function, which is to be integrated many times, is more general than it may at first appear. If a logarithmic term had appeared in the above, we would simply have started our series for [l] at $a_0(p - H)^{-12/7+\epsilon}$, for some suitably chosen small constant ϵ . (See Section 4.)

Suppose we had used a single polynomial of degree N' to approximate l. The resulting approximation, $[s_m]'$, to s_m would be a polynomial of degree m + N'[m/2]. Thus the largest power to which (p - H) would be raised in $[s_m]'$ is m + N'[m/2], whereas the largest power of (p - H) used by our $[s_m]$ is [mN/7]. In the case of air, we used N = 8 and we would have had to use an N' of at least 20, so we compare [8m/7] with m + 20[m/2] in Table 3.2.

TABLE 3.2. The largest power of p - H involved in $[s_m]$ versus that involved in $[s_m]'$.

m	$in \ [s_m]$	$in \ [s_m]'$	
2	2	22	
3	3	23	
4	4	44	
5	5	45	
7	8	67	
10	11	110	
15	17	155	

Thus when p - H takes its largest and smallest values (1.22 and .03125 \cdots in our case), the evaluation of $[s_m]'$ would run into overflow/underflow problems long before the evaluation of our $[s_m]$ does. This saving as well as all the other advantages of our method are due to the facts that

(1) polynomials are "roly-poly," and are not easily fit to an angular function like our l(H) (see Rice [Ric, pp. 15–23]),

(2) functions of the form $\sum_{j=0}^{N-1} a_j(p-H)^{2(j-6)/7}$ behave much like our l(H), especially near H = p, where l is most angular, and

(3) our [l] is in powers of $(p - H)^{2/7}$, beginning at $(p - H)^{-12/7}$, yielding more arbitrary coefficients per degree of (p - H) involved in [l] and hence in $[s_m]$.

If [l] had approximated l to full machine accuracy and if the $c_{m,j}$ in (3.10) were computed in double precision then it would be wasteful to evaluate (3.10) by taking the seventh root of (p - H) and then evaluating (3.10) by a Horner recurrence. Considerable accuracy could be saved by breaking $[s_m]$ into seven terms, each of which is {a polynomial in (p - H)} $\times (p - H)^{n/7}$, for $n = 0, 1, \dots, 6$. The superiority of this method (when the $c_{m,j}$'s are very accurate) can be seen from Eq. (15.2) in Wilkinson [W, p. 50].

4. The Integrals $s_m(H, H_0)$ and Their Approximation for Arbitrary k > 1. For the general case in which k is arbitrary (but > 1), we can apply the methods of Section 3 to show that l has the formal expansion,

(4.1)
$$l(H) = \sum_{j=1}^{\infty} a_j (p-H)^{(k-1)j/k-2}$$
, where $p \equiv H(a_0(2/(k-1))^{1/2})$.

Thus l(H) has a singularity of order (k + 1)/k at p. We can again use this information to find a good form for [l](H). The form chosen must

(1) yield an efficient approximation to l; i.e., it must allow a small maximum error to be obtained by an approximation with few terms (and the coefficients for this approximation should not be large with alternating signs);

(2) yield a simple form for $[s_m]$.

In order to satisfy (2), we must first of all replace (k - 1)/k by a rational approximation, q/r, so that $s_1(H, H_0) = H - H_0$ will be expressible in the form $\sum_j c_{1,j}(p - H)^{j/r}$. As we shall see, it is important to keep r and especially q small. The form

(4.2)
$$[l](H) = \sum_{j=1}^{N} a_j (p-H)^{q(j-2)/r} \text{ for } N \ge 1,$$

will satisfy (1) above, and if no log (p - H) terms enter, we would have

(4.3)
$$[s_m](H, H_0) = \sum_{j=0}^{u_m} c_{m,j} (p - H)^{j/r} \text{ for } N \ge 1,$$

where $u_m = \text{entier}(m/2)qN + r$ and $c_{m,1} = c_{m,2} = \cdots = c_{m,q-1} = 0$. The reason for keeping q and r small is now apparent: the number of terms in (4.3) increases as q and r increase. If log terms would enter, one of the coefficients $c_{m,j}$ would be infinite and this form would fail. In this case we would choose the form

(4.4)
$$[l](H) = \sum_{j=1}^{N} a_j (p-H)^{qj/r-2+\epsilon},$$

where the choice of ϵ is to be explained. Then $[s_m]$ would be

(4.5)
$$[s_m](H, H_0) = \sum_{i=0}^{v_m} \sum_{j=0}^{iqN+r} c_{m,i,j} (p-H)^{j/r+i\epsilon},$$

where $v_m = \text{entier}(m/2)$ and $c_{m,i,1} = c_{m,i,2} = \cdots = c_{m,i,q-1} = 0$. The value of ϵ is chosen so that log terms do not enter, i.e., so that

(4.6)

$$x_{i,j,n} \equiv -1 + \frac{qj+n}{r} + i\epsilon \neq 0,$$

$$x_{i,j,n} + 1 \neq 0,$$
for $i = 1, 2, \cdots$, entier $(m/2) + 1, n = 0, 1, \cdots, iqN + r,$
and $j = 1, \cdots, N,$

for all *m* of interest, say $m = 0, 1, \dots, m$ max. (An irrational value for ϵ will satisfy (4.6) for all *m*, but of course only rational values can be used in current digital computers.) Since terms of the form $(p - H)^{x_{i,j,n}}/x_{i,j,n}$ and $(p - H)^{x_{i,j,n}+1}/(x_{i,j,n}+1)$ will enter, ϵ must be chosen large enough so that $x_{i,j,n}$ and $x_{i,j,n}+1$ are not too small. On the other hand, large values of ϵ will destroy the similarity between (4.4) and (4.2), so ϵ must not be chosen too large.





5. Example. In this section, the approximate solution to an initial value problem is presented in the form of graphs in the hodograph and physical planes. Three similar examples can be found in [B-H-R, pp. 19–40]. The graphs were formed in the following way:

(1) the line $H = H_0$ was specified ($H_0 = -.2$ was used in all four examples), and an Algol procedure was supplied for evaluating $[f](\theta)$, $[g^{(1)}](\theta)$ and their derivatives (these two functions are the initial values for the differential equation);

(2) the coefficients for $[s_m]$, for $m = 0, 1, \dots, 41$ were computed, using the recurrence relations in Theorem 3.1;

(3) the coefficients for $d[s_m]/dH$ were computed from those of $[s_m]$;

(4) three streamlines were traced in the hodograph plane:

$$\Psi(H, \theta) = [\Psi](0, 1.5), \Psi(H, \theta) = [\Psi](.05, 1.5)$$

and

$$\Psi(H, \theta) = [\Psi](.1, 1.5);$$

(5) these streamlines were numerically transformed into the physical plane, using the relations



PHYSICAL PLANE

(See [B-H-R, p. 21] and [S, pp. 28–33] for further details and references concerning this transformation.)

The values of H and θ making up a streamline, $\Psi(H, \theta) = \text{constant}$, were chosen so that $|[\Psi](H, \theta)\text{-constant}| \leq 10^{-5}$. During each calculation of $[\Psi](H, \theta)$, terms in (2.8) were added in until the last term added was $\leq 10^{-6} \times |(\text{the current value of} \text{the sum})|$. An average of six terms (involving $[s_0], [s_1], \dots, [s_{11}]$) of (2.8) was used in computing $[\Psi](H, \theta)$. This example took about 13 minutes on the B5500, and involved about 1300 evaluations of $[\Psi](H, \theta)$. The following initial values were used in this example:

(5.2)
$$f(\theta) = 2.538 \sin r\theta / v(H_0), \ g^{(1)}(\theta) = -2.538 \sin r\theta / (v(H_0)(1 - .2v^2(H_0))^{2.5}),$$

with r = 1.5. The examples in [B-H-R] include r = .8, 1, 1.2. For r = 1, (5.2) gives the initial values for the well-known Ringleb solution [R], 2.538 sin $\theta/v(H)$ (any nonzero constant other than 2.538 is allowable). The example given here resembles flow around a corner; flows of this type are discussed further in [vM, p. 341].

In preparing these examples, v(H) was evaluated by using Newton-Raphson iteration to invert

(5.3)
$$H(v) = \int_{2/(k+1)}^{\tau} \frac{t^{1/(k-1)}}{1-t} dt \quad \text{where } \tau = 1 - \frac{1}{2}(k-1) \left(\frac{v}{a_0}\right)^2$$

(5.4)
$$= \sqrt{\tau}(\tau^2/5 + \tau/3 + 1) - \log\left(\frac{1+\sqrt{\tau}}{1-\sqrt{\tau}}\right) + .251251 \cdots$$

Eq. (5.4) being valid only for k = 1.4. A more efficient method for calculating v(H) is possible if the (approximate) values of $s_m(H, H_0)$ and of $v(H_0)$ are available. And each time $[\Psi_n](H, \theta)$ is evaluated, the values of $[s_m](H, \theta)$, for $m = 0, 1, \dots, 2n + 1$, are available. This method is based on

THEOREM 5.1. Let us define $v_0 = v(H_0)$ and

(5.5)
$$V = \left(1 - \frac{1}{2}(k - 1)\left(\frac{v_0}{a_0}\right)^2\right)^{-1/(k-1)}$$

Then v, v_0, H , and H_0 are related by

(5.6)
$$v(H) = \frac{v_0}{\sum_{j=0}^{\infty} \{s_{2j}(H, H_0) - Vs_{2j+1}(H, H_0)\}}.$$

This result is surprising in that the right side of (5.6) is seen to be independent of H_0 . The relation is most easily derived by equating the Ringleb solution, $\sin \theta/v(H)$, to the solution, as given by (2.7), of the initial value problem, $f(\theta) = \sin \theta/v_0$ and $g^{(1)}(\theta) = -V \sin \theta/v_0$.

Suppose we wish to use (5.6) to calculate v(H) for H in some interval, I. We can use the bounds on $|s_j|$ and $|s_j - [s_j]|$ to be given in Section 6, along with the fact that the denominator in (5.6) has values ranging between

$$\min_{H,H_0 \in I} \frac{v(H_0)}{v(H)} \quad \text{and} \quad \max_{H,H_0 \in I} \frac{v(H_0)}{v(H)},$$

to decide how many terms are needed for the denominator sum in order to make the

truncation error less than or equal the approximation error caused by using $[s_m](H, H_0)$.

6. Error Analysis. Before proceeding with a formal analysis, we present some empirical results. This will allow a more realistic evaluation of the error bounds to be proved. To do this we have used the Ringleb solution,

(6.1)
$$\Psi^{\mathbb{R}}(H,\theta) = \frac{2.538}{v(H)}\sin\theta,$$

of Eq. (2.5) to set up initial value problems for H_0 , $H \in [-1, .22]$. We have then used the program given in [B-H-R] to compute $[\Psi_7^R](H, H_0, \theta)$ for $H, H_0 = -1$, $-.95, \dots, .2$, .22. Figure 6.1 is a graph of the average error, ϵ , versus H_0 , where

(6.2)
$$\epsilon(H_0) \equiv \frac{1}{26} \sum_{j=1}^{26} |\Psi^R(H_j, 1) - [\Psi_7^R](H_j, H_0, 1)|$$

and

 $H_1 = -1, H_2 = -.95, \dots, H_{25} = .20$ and $H_{26} = .22$.

Figure 6.2 is a graph of $|\Psi^R - [\Psi_{7}^R]|$ versus H, for $H_0 = -0.2$. The maximum absolute error tabulated over all these examples was 3.91×10^{-5} , occurring at H = .2, $H_0 = -.95$. The error bound on $|\Psi^R - [\Psi_{7}^R]|$, given by the sum of formulae (6.23) and (6.29), was tabulated for $H_0 = -1, -.95, \cdots, .05$ and $H = -1, -.95, \cdots, .2$, .22 (the omission of $H_0 = .1, .15, .2, .22$ will be explained shortly). The upper curves in Figs. 6.1 and 6.2 are the corresponding graphs for this error bound. The maximum value tabulated for this bound was 1.2×10^{-3} , occurring at $H = .22, H_0 = -1.0$. It is difficult to maximize this bound, as a function of H and H_0 . However, a somewhat weaker bound, given by (6.37) + (6.38), can be maximized easily, yielding an upper bound (for all $H_0 \in [-1, .06593 \cdots]$ and $H \in [-1, .22]$) on the error in our approximate Ringleb solution of 3.3×10^{-3} .



Fig. 6.1

It should be pointed out that the bounds of this section depend on $\delta \equiv \max_{H \in [\alpha,\beta]} |l(H) - [l](H)|.$

To get the values of the bounds discussed above, it was necessary to use (3.6) to

get a value for δ . As mentioned earlier, (3.6) is not a mathematically established relation, so when we set $\delta = 4.1 \times 10^{-5}$, we do not get mathematically established bounds. But we do get quite believable bounds (because (3.6) is quite believable).



These calculations were done only for $\theta = 1$ radian since the simple form of Ψ^{R} and the fact that the error in $[f^{(2j)}]$ and $[g^{(2j+1)}]$ is very small in this case, make the *relative* error given by the formulae of this section essentially independent of θ .

Let us proceed with a formal error analysis. The error involved in our computation draws from three sources:

(1) truncation—we have truncated the infinite series (2.7) for Ψ to yield Ψ_n ;

(2) function approximation—we have permitted the use of [l], $[f^{(2i)}]$ and $[g^{(2i+1)}]$, for $j = 0, 1, \dots, n$, to yield $[\Psi_n]$; and

(3) roundoff—computations are done in fixed, finite precision arithmetic.

Errors of types (2) and (3) can be confused easily: type (2) errors are due to the fact that the formulae used to calculate certain functions would not give exact values, even if exact arithmetic were used; type (3) errors are due to the inexactness of computer arithmetic. Confusion may arise when the inexact formulae are correct to within the roundoff error of the inexact arithmetic.

Roundoff error has been no problem in our work, partly because we are using 10 digits for our essentially 5-digit calculations. We shall not consider roundoff error here. The following analysis provides *absolute* bounds, as functions of H, H_0 and θ , for the truncation and function approximation errors. A series of five lemmas is required. The first three lemmas present rough bounds based on (2.9), itself a rather rough bound on $|s_m|$. The derivation of these rough bounds utilizes only one property of l, namely that for $H \in [\alpha, \beta]$, $|l(H)| \leq c^2$. In this paper, we deal with $[\alpha, \beta] \subseteq [-1, .22]$, for which $c^2 \leq 62.47$. When evaluating our bounds for particular H and H_0 , we of course choose $[\alpha, \beta] = [H_0, H]$, and use a corresponding c.

Let a be defined by

(6.3)
$$l(a) = -1$$
.

(For k = 1.4, we have $a = .0659262218 \cdots$.) When $H_0 << a << H$ or H <<

 $a < < H_0$, the first bounds are poor. Lemmas 6.4 and 6.5 give considerably improved bounds, valid for $H_0 \leq a \leq H$. In the Ringleb computation considered, these new bounds were as much as 10¹⁰ better than the old bounds. The case $H \leq a \leq H_0$ probably can be treated similarly, but this will not be done here. (This is why the cases $H_0 = .1, .15, .2, .22$ were omitted from the bound calculations summarized in Figs. 6.1 and 6.2.) The improved bounds depend on one further property of l, namely that $|l(H)| \leq 1$ for $H \in [\alpha, a]$ with $\alpha \leq a$ (and for any k > 1).

In order to present simple a priori bounds, we assume that, for fixed θ , $f^{(2j)}(\theta)$ and $q^{(2j+1)}(\theta)$ grow (with j) no faster than geometrically. However, the derivatives of analytic functions can grow much faster than this. (If $h(\theta)$ is analytic, then by Cauchy's formula, $|h^{(j)}(\theta)| \leq \max |h(\theta)| j! r^{-j-1}$, where r is the minimum distance of θ from the boundary of some domain within which h is analytic; the maximum of $|h(\theta)|$ is to be taken over the same domain from which r is computed.) The bound on the approximation error also involves terms which must bound the error caused by $[f^{(2j)}]$ and $[q^{(2j+1)}]$ for $j \leq n$. If these errors can be assumed negligible (or if a bound can be found), then an a posteriori bound on the error due to function approximation can be computed, while the approximate stream function, $[\Psi]$, is being computed, without any assumptions about the growth of $f^{(2j)}$ and $g^{(2j+1)}$; the actual values of $[f^{(2_j)}](\theta)$ and $[g^{(2_j+1)}](\theta)$ could be used in the bounds. This is not possible for the truncation error; we must have definite knowledge of the growth of $f^{(2i)}$ and $g^{(2i+1)}$, as $j \to \infty$, in order to bound this error. And a bound on the function approximation error is of no value without a bound on the truncation error. The usual heuristic solution to this problem consists of letting the program determine when to truncate the series for Ψ dynamically, on the basis of the size of the last term computed; when the last term is small relative to the current value of the series, the truncation error would be assumed negligible. (The program given in [B-H-R] allows the user to decide whether a fixed number of terms or the heuristic stopping criterion is to be used.)

In the following, we assume that c > 0, and we let T_n and A_n denote the truncation and function approximation errors involved in (2.8), respectively, so that

(6.4)
$$T_n(H, H_0, \theta) \equiv \Psi(H, \theta) - \Psi_n(H, H_0, \theta) ,$$

(6.5)
$$A_n(H, H_0, \theta) \equiv \Psi_n(H, H_0, \theta) - [\Psi_n](H, H_0, \theta).$$

The proofs of the following lemmas may be found in [B-H-R].

LEMMA 6.1. Let θ be fixed. Suppose there exist constants r_f , r_g , B_f and B_g for which

(6.6)
$$|f^{(2j)}(\theta)| \leq r_f^{2j} B_f$$
, $|g^{(2j+1)}(\theta)| \leq r_g^{2j+1} B_g$ for $j \geq n+1$.

Let an upper bound function, U_n , be defined by

(6.7)
$$U_n(h,x) \equiv B_h \frac{(r_h x)^n}{n!} \cosh(r_h x) ,$$

where h can be f or g. Then we have

(6.8)
$$|T_n(H, H_0, \theta)| \leq U_{2n+2}(f, c|H - H_0|) + \frac{1}{c} U_{2n+3}(g, c|H - H_0|)$$

for all $H, H_0 \in [\alpha, \beta]$.

Let us define

(6.9)
$$S_m(H, H_0) \equiv \frac{\delta_m^{-1}}{m!} (c|H - H_0|)^m,$$

(6.10)
$$E_m(H, H_0) = s_m(H, H_0) - [s_m](H, H_0),$$

(6.11)
$$\delta \equiv \max_{H \in [\alpha,\beta]} |l(H) - [l](H)|,$$

where $\delta_m = 1$ if m is even and $\delta_m = c$ if m is odd. LEMMA 6.2. We have

(6.12)
$$|E_m(H, H_0)| \leq \frac{\delta}{c^2} \operatorname{entier}\left(\frac{m}{2}\right) S_m(H, H_0) (1 + \delta c^{-2})^{m/2} \text{ for } m \geq 0.$$

LEMMA 6.3. Let θ be fixed, and let constants C_f , D_f , C_g , D_g , c_f , c_g , d_f and d_g satisfy

(6.13)
$$C_{f}c_{f}^{2j} \ge |f^{(2j)}|, \quad C_{g}c_{g}^{2j+1} \ge |g^{(2j+1)}|,$$

$$(6.14) D_{g} d_{g}^{2j} \ge |f^{(2j)} - [f^{(2j)}]|, D_{g} d_{g}^{2j+1} \ge |g^{(2j+1)} - [g^{(2j+1)}]|,$$

for $j = 0, 1, \dots, n$. Let us define bounding functions, F and G, by

(6.15)
$$F(K, x, y) \equiv \frac{\delta}{2K} \left(C_f x \sinh x + D_f y \sinh y \right) + D_f \cosh y ,$$

(6.16)
$$G(K, x, y) \equiv \frac{\delta}{2K} \left(C_g x (\cosh x - 1) + D_g y (\cosh y - 1) \right) + D_g \sinh y \, .$$

Then we have, with $z = (1 + \delta c^{-2})^{1/2} |H - H_0|c$,

(6.17)
$$|A_n(H, H_0, \theta)| \leq F(c^2, c_f z, d_f z) + \frac{1}{c} G(c^2, c_g z, d_g z),$$

independent of n.

The above bounds on T_n and A_n are reasonable as long as $[\alpha, \beta]$ is such that c remains small. But as $\beta \to p$ we have $c \to \infty$. Our bounds can be weak because the constant c multiplies the whole of $|H - H_0|$ in our bound of (2.9):

(6.18)
$$|s_m(H, H_0)| \leq (c|H - H_0|)^m \delta_m^{-1}/m!$$

If $H_0 << a << H$, then c and $|H - H_0|$ are large. It does not seem fair that, in this case, c should multiply all of $|H - H_0|$ since c is only needed to bound l in [a, H]; a bound of unity suffices in $[H_0, a]$. Thus we may expect to be able to replace $c|H - H_0|$ by $c(H - a) + a - H_0$ in this case. Indeed, this can be done if the factor of δ_m^{-1} is removed, as can be proved from the following stronger result. With h =H - a and $h_0 = H_0 - a$, let us define

(6.19)
$$S_m^*(H, H_0) \equiv .5\left(1 + \frac{1}{c}\right)\frac{(ch - h_0)^m}{m!} + .5\left(1 - \frac{1}{c}\right)\frac{(-ch - h_0)^m}{m!}$$

for $m \ge 0$.

LEMMA 6.4. We have

(6.20)
$$|s_m(H,H_0)| \leq S_m^*(H,H_0) \quad \text{for } H_0 \leq a \leq H,$$

with equality holding for m = 0, 1. Further, this bound holds if a is replaced by any number between H_0 and a. If a is replaced by H_0 or c = 1, then (6.20) reduces to (6.18). Also, we have

$$(6.21) \quad S_m(H, H_0) > S_m^*(H, H_0) \quad for H_0 < a < H \text{ and } m \ge 2,$$

(6.22)
$$S_m(H, H_0) = S_m^*(H, H_0)$$
 for $H_0 = a \leq H$ or $H_0 \leq a = H$ or $m = 0, 1$.

The case $H \leq a \leq H_0$ probably can be dealt with in a similar manner, but this will not be pursued here. The bound on T_n corresponding to this new bound is

$$|T_{n}(H, H_{0}, \theta)| \leq .5\left(1 + \frac{1}{c}\right) \{U_{2n+2}(f, ch - h_{0}) + U_{2n+3}(g, ch - h_{0})\}$$

$$(6.23) + .5\left(1 - \frac{1}{c}\right) \{U_{2n+2}(f, -ch - h_{0}) + U_{2n+3}(g, -ch - h_{0})\}$$

$$for H_{0} - a = h_{0} \leq 0 \leq h = H - a.$$

To get a new bound on E_m and A_n we present the following generalization of (6.12).

LEMMA 6.5. If $E_m(H, H_0)$ and $S_m^*(H, H_0)$ are defined as in (6.10) and (6.19) then

where $S_{-1}^* \equiv 0$, and $\sigma(m) = 0$ if m is even and $\sigma(m) = 1$ if m is odd. Further, this holds if a is replaced by any number in $[H_0, a]$; if a is replaced by H_0 and $(1 + \delta)^{m/2}$ by $(1 + \delta c^{-2})^{m/2}$, or if c = 1 then this reduces to (6.12).

Various weaker, but simpler, bounds can be proved, two of the simplest (and weakest) being $\delta(1 + \delta)^{m/2}$ entire $(m/2) S_m^*(H, H_0)$ and

$$\delta$$
 entier $(m/2)$ $((ch - h_0) (1 + \delta)^{1/2})^m/m!$

The new bound (6.24) on E_m provides the following bound on A_n : let bounding functions F_i and G_i be defined by

(6.26)

$$F_{1}(K, x, y) \equiv \left(1 + \frac{1}{c}\right)(x + b(x + y))K \sinh(Kx) + \left(1 - \frac{1}{c}\right)(y + b(x + y))K \sinh(Ky),$$

$$F_{2}(x, y) \equiv \frac{\delta}{4c^{2}} \{C_{f}F_{1}(c_{f}, x, y) + D_{f}F_{1}(d_{f}, x, y)\} + \frac{D}{2} \left\{\left(1 + \frac{1}{c}\right)\cosh(d_{f}x) + \left(1 - \frac{1}{c}\right)\cosh(d_{f}y)\right\},$$

(6.27)

$$G_{1}(K, x, y) \equiv \left(1 + \frac{1}{c}\right)\left(x + \frac{b}{2}(x + y)\right)K(\cosh(Kx) - 1) + \left(1 - \frac{1}{c}\right)\left(y + \frac{b}{2}(x + y)\right)K(\cosh(Ky) - 1) - b(x + y)(\cosh(K(x - y)/2) - 1),$$

 $G_{2}(x, y) \equiv \frac{\delta}{4c^{2}} \{ C_{g}G_{1}(c_{g}, x, y) + D_{g}G_{1}(d_{g}, x, y) \}$

(6.28)

$$+ rac{D_g}{2} \left\{ \left(1 + rac{1}{c}
ight) \sinh \left(d_g x
ight) + \left(1 - rac{1}{c}
ight) \sinh \left(d_g y
ight)
ight\},$$

where $b = c^2 - 1$. Then it follows that

(6.29)
$$|A_n(H, H_0, \theta)| \leq F_2(x, y) + G_2(x, y)$$

where

(6.30)
$$x = (ch - h_0)(1 + \delta)^{1/2}$$
 and $y = (-ch - h_0)(1 + \delta)^{1/2}$

Our new bounds, (6.23) and (6.29), reduce to the old bounds when either c = 1 or a is replaced by H_0 , $(1 + \delta)^{m/2}$ by $(1 + \delta c^{-2})^{m/2}$ and, if $H_0 > H$, then H and H_0 are interchanged. For this reason, a program for calculating these bounds need be written only for (6.23) and (6.29); for the cases H < a or $H_0 > a$, the old bounds can be derived by the replacement just described. For the Ringleb computation, all growth constants are 1, and

(6.31)
$$C_f = B_f = |2.538 \sin(1)/v_0|,$$

(6.32)
$$C_g = B_g = \left| \frac{2.538 \sin(1)}{v_0 (1 - .2v_0^2)^{2.5}} \right|,$$

(6.33)
$$D_h = 10^{-9} B_h \text{ for } h = f, g$$

$$\delta = 4.10533 \times 10^{-5}.$$

The bounds

(6.35)
$$|s_m(H, H_0)| \leq \frac{(ch - h_0)^m}{m!}$$
 for $H_0 \leq a \leq H$,
(6.36) $|E_m(H, H_0)| \leq \delta$ entier $\left(\frac{m}{2}\right) \frac{((ch - h_0)(1 + \delta)^{1/2})^m}{m!}$ for $H_0 \leq a \leq H$,

m!

can be used to derive simpler bounds on
$$A_n$$
 and T_n :

$$(6.37) |A_n(H, H_0)| \leq F(1, c_f z, d_f z) + G(1, c_g z, d_g z),$$

$$(6.38) |T_n(H, H_0, \theta)| \leq U_{2n+2}(f, ch - h_0) + U_{2n+3}(g, ch - h_0),$$

where $z = (ch - h_0) (1 + \delta)^{1/2}$ and F and G are given by (6.15) and (6.16). As $ch - h_0$ increases and H_0 decreases, these bounds increase. Thus they attain their maxima when $H = \beta$ and $H_0 = \alpha$. For the Ringleb computation described above, this implies

 $|T_7| + |A_7| \le 3.3 \times 10^{-3}$ for $H \in [-1, .22]$ and $H_0 \in [-1, a]$, (6.39)

the bound being calculated at H = .22 and $H_0 = -1$. The disadvantage of these simpler bounds is that, when a is replaced by H_0 , they do not reduce to our old bounds; a factor of c^2 is lost. Thus, as $H_0 \rightarrow a$ from below, while H > a, these bounds will become several orders of magnitude worse than our more complex bounds. (If β were closer to p, then c^2 would be even larger and this loss would be more drastic.)

Bell Telephone Laboratories Murray Hill, New Jersey 07974

B.1. S. BERGMAN, Integral Operators in the Theory of Linear Partial Differential Equations, Ergebnisse der Mathematik und ihrer Grenzgebiete, Heft 23, Springer-Verlag, Berlin, 1961. MR 25 #5277.

B.2. S. BERGMAN, "On representation of stream functions of subsonic and supersonic flows of compressible fluids," J. Rational Mech. Anal., v. 4, 1955, pp. 883-905. MR 17, 549.
B.3. S. BERGMAN, "Two-dimensional subsonic flows of a compressible fluid and their singularities," Trans. Amer. Math. Soc., v. 62, 1947, pp. 452-498. MR 10, 162.
B.4. S. BERGMAN, "Applications of the kernel function for the computation of flows of compressible fluids," (To groups).

pressible fluids." (To appear.)

B-H-K. S. BERGMAN, J. G. HERRIOT & T. G. KURTZ, "On numerical calculation of transonic

B-HTER. S. BERGMAN, J. G. HERRIOT & P. 13-27. MR 36 #7379.
 B-H-R. S. BERGMAN, J. G. HERRIOT & P. L. RICHMAN, On Computation of Flow Patterns of Compressible Fluids in the Transonic Region, Technical Report No. CS70, Computer Science De-

Compression Future Future for the first field of the potential function of a two dimensional formation of the potential function of the potential functin of the potential function of the potential function of the

Chebyshev's system of functions." (To appear.)
L. G. S. S. LUDFORD, "The behavior at infinity of the potential function of a two dimensional subsonic flow," J. Mathematical Phys., v. 30, 1950, pp. 117-130. MR 13, 399.
Ric. J. R. RICE, The Approximation of Functions. Vol. I. Linear Theory, Addison-Wesley, Reading, Mass., 1964. MR 29 #3795.
R. F. RINGLEB, "Exakte Lösungen der Differentialgleichungen einer adiabatischen Gasströmung," Z. Angew. Math. Mech., v. 20, 1940, pp. 185-198. MR 2, 169.
S. J. M. STARK, "Applications of Bergman's integral operators to transonic flows," Internat. J. Non-linear Mech., v. 1, 1966, pp. 17-34.
vM. R. von MISES, Mathematical Theory of Compressible Fluid Flow, Academic Press, New York, 1958. MR 20 #1504.
W. J. H. WILKINSON, Rounding Errors in Algebraic Processes, Prentice-Hall, Englewood Cliffs, N. J., 1963. MR 28 #4661.

372