Hansen

APPLIED MATHEMATICS AND STATISTICS LABORATORIES

STANFORD UNIVERSITY

CALIFORNIA

INTERVAL ARITHMETIC AND AUTOMATIC ERROR ANALYSIS IN DIGITAL COMPUTING

BY

R. E. MOORE

TECHNICAL REPORT NO. 25

NOVEMBER 15, 1962

PREPARED UNDER CONTRACT Nonr-225(37)

(NR-044-211)

FOR

OFFICE OF NAVAL RESEARCH

INTERVAL ARITHMETIC AND AUTOMATIC ERROR ANALYSIS

IN DIGITAL COMPUTING

by

R. E. Moore

TECHNICAL REPORT NO. 25

November 15, 1962

PREPARED UNDER CONTRACT Nonr-225(37)

(NR-044-211)

OFFICE OF NAVAL RESEARCH

APPLIED MATHEMATICS AND STATISTICS LABORATORIES
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

PREFACE

Very lengthy numerical computations were performed until recent years, just as the shortest computations, by hand or with the aid of desk calculators.

In an hour's time a modern high-speed stored-program digital computer can perform arithmetic computations which would take a "hand computer" equipped with a desk calculator five years to do.

In setting out on a five year computing project, a hand computer would be justifiably (and very likely gravely) concerned over the extent to which errors were going to accumulate — not mistakes, which he will catch by various checks on his work — but errors due to rounding and the replacement of integrals by finite sum approximations, etc. In order to be able to guarantee the accuracy of his final results to a preassigned number of significant figures, a substantial analytical effort will be required in choosing certain parameters such as how many terms of a series are needed, in arranging the order of the work, in deciding how many places should be carried, etc.

A complete a priori error analysis for an extensive numerical computation can become a formidable task if it is to answer in advance and in every detail all questions of the accuracy of approximations to be made during the computation and their effect on the accuracy of the final results.

While he will do some of the analysis in advance, the hand computer will leave a fair amount of it to be done during the course of the numerical work — taking advantage of partial results occurring in the specific computation at hand to sharpen bounds and simplify the analysis.

In order to take full advantage of the great speed of the automatic stored-program digital computer it is obviously desirable to mechanize as much as possible of the error analysis required for a computation so that it can be carried out by the machine itself.

It is the intent of the present study to investigate an approach, based on <u>interval arithmetic</u>, by which a stored-program digital computing machine can be made to produce rigorous error bounds during the course of a computation — error bounds which, at least for some types of computation, will be sharp enough and easily enough computed so that <u>automatic error analysis</u> by the machine can replace the lengthy a priori analysis otherwise required in advance of a computation if accuracy is to be guaranteed.

The conversion of an automatic error analysis scheme to an automatic error control scheme is relatively easy (see Reference [7], for example). In the end we can finally give the computing machine instructions similar to those we would give a hand computer. For example: "compute an approximation to the solution of the following problem ... , accurate to ... decimal places, unless it would take longer than ... hours, in which case get the most accuracy you can in that time."

Machine programs of this type have, in fact, been written (see [6], [7]); however, the error bounds upon which those error control schemes are based (in [6], [7]) are rigorous only if certain conditions are met which are usually very difficult to verify. In the present study the methods based on interval arithmetic are much more general and yield rigorous error bounds with only trivial conditions to be met. Even these conditions can be tested by the machine. At one point, for example, we require an interval which contains no poles of a given rational function. We can choose some interval and ask the question: Does the given rational function have any poles in the chosen interval? Either the machine will succeed in computing a finite interval containing the range of values of the function for the given interval of arguments or it will halt on an attempted division by an interval containing zero. In the former case the question is answered in the negative. In the latter case the question remains unanswered; but we can try again with a smaller interval and eventually find an interval containing no poles.

A rounded version of interval arithmetic provides a means for the automatic computation by the machine of intervals containing the infinite precision result of any computation of finite length with real numbers.

ACKNOWLEDGMENTS

This work was supported financially by the Lockheed Aircraft Corporation. Machine programming and numerical tests were carried out at Lockheed in Palo Alto, California on the 1103AF and 7090 computers, principally by D. Thoe, with assistance from A. Dempster, A. Podvin and A. J. Cook. Professors W. Strother and C. T. Yang, while acting as Lockheed consultants, contributed both in actual work and in interest. (See references [13], [14].) Support and encouragement at Lockheed was given by C. E. Duncan, B. D. Rudin, R. J. Dickson, and others. From its earliest stages in 1957 Professor G. E. Forsythe of Stanford University encouraged the work and made many valuable suggestions. The typing and reproduction was done under Office of Naval Research contract Nonr-225(37), at Stanford University.

TABLE OF CONTENTS

INTERVAL ARITHMETIC AND AUTOMATIC ERROR ANALYSIS

IN DIGITAL COMPUTING

1. Introduction.

Computations by digital computing machines consist of finite sequences of rounded arithmetic operations. Such computations are most often performed as approximations to finite or infinite sequences of exact arithmetic operations with real numbers.

A digital computation and an analysis of its error as an approximation are necessarily viewed as separate processes when one is dealing with real numbers. On the other hand, in the present study an interval arithmetic is described which forms a basis for a concomitant analysis of error in a digital computation. In this system computations with intervals are performed and intervals are so produced to contain, by construction, the exact numerical solutions sought. Hence, an approximation and an error bound are obtained at the same time; choosing say the midpoint of the interval as the approximation, the radius of the interval becomes the error bound.

Earlier discussion of interval arithmetic can be found in Dwyer [3], [4], Fischer [5], Sunaga [20], Moore [12], and Moore and Yang [13].

The theory of real valued functions of intervals has been developed in connection with real integration theory: [1], [2], [16], and [17]. An integration theory for continuous interval valued functions has been considered in [13] and [14]. An algebraic system more general than interval arithmetic appeared in [21]. However, the emphasis there is set theoretical rather than computational.

In the present study the principal results will be convergence theorems for interval computations of: the range of values of a rational function (Th. 4.1), the definite integral of a rational function (Th. 5.1, Th. 5.4), solutions to ordinary differential equations in rational functions (Th. 6.8, Th. 6.16). The methods will be illustrated by examples. Theorem 4.1 and Th. 5.1 are generalizations of results obtained in [13].

We first develop the theory for exact or infinite precision interval arithmetic in order to obtain finite computations which bound the results of infinite sequences of computations. Finally, for actual machine computations, rounded interval arithmetic is introduced so that round-off errors are also taken into account. In fact, in this way machine programs using rounded interval arithmetic and the computational schemes developed in the following sections will produce with finite machine computations rigorous upper and lower bounds to the solutions of various analytical problems.

The widths of the bounding intervals can be made as small as one pleases at the cost of the amount of computation required.

Numerical examples of actual machine work indicate that usefully narrow bounds are obtainable, at least for some types of problems, in reasonable computing time even with the straightforward first version of the theory developed here. It is not unlikely that many improvements in efficiency of the techniques can be made.

2. Interval Arithmetic.

Denote by $[a,b]$ the closed interval of real numbers x such that $a \leq x \leq b$. Let $\mathcal{I}$ be the set of all such intervals and let $\mathcal{I}^*$ be the set of intervals which do not contain the real number 0.

Arithmetic operations in $\mathcal{I}$ are defined as follows, with I, J in $\mathcal{I}$ and K in $\mathcal{I}^*$:

$$
(2.1) \quad \left\{ \begin{aligned}
I + J &= \{x + y \mid x \in I,\ y \in J\} \equiv \text{the set of real numbers } t \text{ such that } t = x + y \text{ for some } x \in I \text{ and some } y \in J, \\
I - J &= \{x - y \mid x \in I,\ y \in J\}, \\
IJ &= \{xy \mid x \in I,\ y \in J\}, \\
I \div K &= I/K = \{x/y \mid x \in I,\ y \in K\}.
\end{aligned} \right.
$$

We will use the usual "power" notation for self-products, i.e., $X^2 \equiv XX$, etc.

The real arithmetic operations are continuous and so map compact connected sets and, in particular, rectangles onto compact connected sets, that is onto closed real intervals. In fact, for $[a,b]$, $[c,d] \in \mathcal{I}$ we find that

$$
(2.2) \quad \left\{ \begin{aligned}
[a,b] + [c,d] &= [a + c,\ b + d] \\
[a,b] - [c,d] &= [a - d,\ b - c] \\
[a,b]\,[c,d] &= [\min(ac, ad, bc, bd),\ \max(ac, ad, bc, bd)]
\end{aligned} \right.
$$

and for $[a,b] \in \mathcal{I}$, $[e,f] \in \mathcal{I}^*$ we have

$$[a,b] \,/\, [e,f] = [a,b]\,[1/f,\, 1/e] \,.$$

By the equality of two intervals $[a,b] = [c,d]$ we mean, of course, that $a = c$ and $b = d$. Alternatively, $I = J$ iff*/ $I \subset J$ and $J \subset I$, regarding I and J as sets of real numbers.

For intervals of the form $[a,a]$ we have

$$[a,a] + [b,b] = [a + b,\, a + b]$$

$$[a,a] - [b,b] = [a - b,\, a - b]$$

$$[a,a]\,[b,b] = [ab,\, ab]$$

and if $b \neq 0$

$$[a,a] \,/\, [b,b] = [a/b,\, a/b] \,,$$

so that interval arithmetic includes real arithmetic, identifying the interval $[a,a]$ with the real number a. We make this identification henceforth and treat the real number field as a subsystem of the interval number system.

Associativity and commutativity for interval addition and interval multiplication are immediate consequences of the definitions (2.1). The real numbers 0, 1 (i.e., the intervals [0,0] and [1,1]) serve as additive and multiplicative identities respectively:

---

*/ If and only if.

$$0 + I = I + 0 = I$$

$$1I = I1 = I \ .$$

However, inverses do not in general exist and the distributive law does not always hold. Indeed,

$$[a,b] - [c,d] = [a - d, b - c] = 0 , \qquad \text{iff} \quad a = d , b = c .$$

Since $a \leq b$, $c \leq d$, this means that

$$[a,b] - [c,d] = 0 , \qquad \text{iff} \quad a = b = c = d .$$

Thus the only intervals having additive inverses are the real numbers. Similarly, $[a,b]\,[c,d] = 1$ iff $[a,b] = x$, $[c,d] = x^{-1}$ for a real number x, i.e., $x = a = b = c^{-1} = d^{-1}$.

The distributive law fails, since*/

$$[1,2](1 - 1) = [1,2]0 = 0 ,$$

whereas

$$[1,2](1) + [1,2](-1) = [1,2] + [-2,-1]$$

$$= [-1,1] \neq 0 .$$

---

*/ Rounded arithmetic with real numbers is also not distributive; in fact, it is not even associative; see [15], p. 1031, and also Section 7 below.

Nevertheless, we do have the following law: for $I, J, K \in \mathcal{I}$,

$$I(J + K) \subset IJ + IK \quad . \tag{2.3}$$

This relation, which we will call subdistributivity, follows easily from the definitions (2.1).

Some special cases in which $IJ + IK = I(J+K)$ (i.e., distributivity holds) are useful. In particular, if t is real, then

$$t(J + K) = tJ + tK \tag{2.4}$$

(This follows immediately from (2.1)); if $JK \geq 0$ (that is, if $x \in JK \Rightarrow x \geq 0$), then

$$I(J + K) = IJ + IK \ . \tag{2.5}$$

To show (2.5), it is sufficient to show $IJ + IK \subset I(J + K)$ because of (2.3). Suppose $x \in IJ + IK$, then

$$x = t_1 y + t_2 z \ ,$$

for some

$$t_1 \ , t_2 \in I \ , \quad y \in J \ , \quad z \in K \ .$$

By hypothesis $yz \geq 0$, so we have $y + z = 0$ iff $y = z = 0$. If $y = z = 0$, then x is clearly in $I(J + K)$. Otherwise $y + z \neq 0$, and choosing

$$t = t_1 \frac{y}{y+z} + t_2 \frac{z}{y+z} \in I ,$$

we have

$$x = t(y+z) \in I(J+K) .$$

In order to study finite sequences of interval arithmetic operations, the following relations are useful.

The elements of $\mathcal{I}$ are partially ordered by set inclusion. In fact, $[a,b] \subset [c,d]$ iff $c \leq a \leq b \leq d$.

The arithmetic operations in $\mathcal{I}$ are <u>inclusion monotonic</u>, i.e., if $I, J, K, L \in \mathcal{I}$, $I \subset K$, $J \subset L$, then

$$(2.6) \qquad \begin{cases} I + J \subset K + L , \\ I - J \subset K - L , \\ IJ \subset KL , \\ I/J \subset K/L , \qquad (0 \notin L) . \end{cases}$$

These relations follow immediately from the definitions (2.1). Together they have the important consequence that if $F(X_1, X_2, \ldots, X_n)$ is a <u>rational</u> expression in the interval variables $X_1, X_2, \ldots, X_n$, i.e., a finite combination of $X_1, \ldots, X_n$ and a finite set of constant intervals in an expression with interval arithmetic operations, then

$$X_1' \subset X_1 , \quad X_2' \subset X_2 , \ldots , X_n' \subset X_n$$

implies

$$(2.6)' \qquad F(X_1', X_2', \ldots , X_n') \subset F(X_1, X_2, \ldots , X_n) .$$

An interesting consequence of subdistributivity (see (2.3), together with (2.6)') is the fact that a "nested" interval polynomial is "contained in" the corresponding power polynomials, i.e.,

$$A_o + X(A_1 + X(A_2 + \cdots + X(A_n)) \cdots) \subset A_o + A_1 X + \cdots + A_n X^n .$$

Notice that a rational interval form is not usually representable by a quotient of two polynomials. We cannot write

$$X + \frac{1}{X} = \frac{X^2 + 1}{X} ,$$

for example, since as we have already noted $X/X \neq 1$. Indeed, let

$$F_1(X) = X + \frac{1}{X} , \qquad F_2(X) = \frac{X^2 + 1}{X} ;$$

then

$$F_1([1,2]) = [1,2] + \frac{1}{[1,2]} = [1,2] + \left[\frac{1}{2}, 1\right] = \left[\frac{3}{2}, 3\right] ,$$

whereas

$$F_2([1,2]) = \frac{[1,2]^2 + 1}{[1,2]} = \frac{[1,4] + 1}{[1,2]} = \frac{[2,5]}{[1,2]} = [1,5] .$$

The two expressions $X + \frac{1}{X}$, $\frac{X^2 + 1}{X}$ define distinct interval valued functions F_1, F_2, say, on the domain $X > 0$; i.e., $D = \{I \mid x \in I \Rightarrow x > 0\}$. These functions do, of course, have the same real restriction:*/ the real rational function given by $f(x) = \frac{x^2+1}{x}$ $(x > 0)$, since for real x, $x + \frac{1}{x} = \frac{x^2 + 1}{x}$. Notice that, if a rational interval expression can be evaluated for a set of values $A_1 , \dots , A_n$ of its variables $X_1 , \dots , X_n$ (which is to say that no division by an interval containing zero occurs), then it can also be evaluated for any set of values $A_1' , \dots , A_n'$ of the variables $X_1 , \dots , X_n$, such that $A_1' \subset A_1 , \dots , A_n' \subset A_n$. We will call a set D of n-tuples of intervals a regular domain if

$$(A_1, A_2, \dots , A_n) \in D \quad \text{and} \quad A_1' \subset A_1 , \dots , A_n' \subset A_n$$

together imply

$$(A_1', \dots , A_n') \in D .$$

Denote by $\mathcal{J}_A$ the set $\{A' \mid A' \subset A\}$, then a regular domain is a union of sets of the form $\mathcal{J}_{A_1} \otimes \mathcal{J}_{A_2} \otimes \cdots \otimes \mathcal{J}_{A_n}$. We use $\otimes$ for the Cartesian product.

*/ By the real restriction of a function on intervals we mean the function restricted to the subdomain consisting of the special intervals of the form $[x,x]$, i.e., real numbers. The value of the function may not be real even at a point of this subdomain. For example, $F([a,b]) = [0,1]$ for $a \leq b$, has real restriction $f(x) = F([x,x]) = [0,1]$.

We define a <u>rational</u> interval function F in n variables $X_1, \dots, X_n$ to be a mapping $f : D \to \mathcal{I}$ with regular domain $D \subset \mathcal{I}^n$ such that there is a rational interval expression in the variables $X_1, \dots, X_n$ which represents the function F. We now have the following. If f is the real restriction of a rational interval function F with domain D, then

$$(2.7) \qquad \bigcup f(x_1, x_2, \dots, x_n) \subset F(X_1, \dots, X_n),$$

with the union taken over $x_i \in X_i$, $i = 1, 2, \dots, n$.

The computational significance of this result is illustrated by the following application to the problem of bounding zeros of rational functions.

Suppose f is a rational function of a real variable x with real coefficients. Let F be a rational interval function of X with real coefficients such that the real restriction of F is f, i.e.,

$$F(x) = f(x), \qquad \text{for } x \text{ real}.$$

Furthermore, let F' be a rational interval function whose real restriction is $f' = \frac{df}{dx}$. If $A = [a,b]$ is an interval, such that

$$f(a) = F(a) < 0 < F(b) = f(b)$$

and

$$F'(A) \in \mathcal{I}^*,$$

then f' is bounded and has the same sign for every $x \in A$ and f has a unique zero in A. By the mean value theorem if x, $x + h \in A$, then $f(x + h) = f(x) + f'(x + \theta h)h$ for some $\theta \in [0,1]$.

Suppose x is the zero of f in A, and y is an arbitrary point in A, then

$$x = y - \frac{f(y)}{f'(x + \theta(y - x))} \qquad \text{for some} \quad \theta \in [0,1] \ .$$

By (2.7), we have

$$x \in y - \frac{F(y)}{F'(A)} \ .$$

If A is not too wide, then the interval

$$y - \frac{F(y)}{F'(A)}$$

will be properly contained in A; and, in fact, the sequence of intervals, each containing the zero of f obtained by iterating the process, will converge to an interval of zero width which must therefore be the zero of f. We omit the easy proof of this assertion and instead give an example to show the mechanism of the procedure.

Denote by the mX the midpoint of the interval X; if $X = [a,b]$, then $mX = \frac{a + b}{2}$. Let $A_o = A$, $y_o = mA$, and for $i = 1, 2, \ldots$, define

$$A_{i+1} = y_i - \frac{F(y_i)}{F'(A_i)} ,$$

$$y_{i+1} = mA_{i+1} .$$

The midpoint y_i was chosen for symmetry. We could just as well use any point in A_i. The procedure is nothing more than the well-known "Newton's method" for finding zeros of a function modified to yield the bounding intervals, A_i. The width of the interval A_{i+1} is roughly proportional to the square of the width of A_i.

For a numerical example, take

$$f(x) = x^2 - 2 , \qquad A = A_o = [1,2] ,$$

$$F(X) = X^2 - 2 , \quad F'(X) = 2X .$$

Then the procedure gives

$$A_{i+1} = y_i + \left(1 - \frac{1}{2} y_i^2\right) \frac{1}{A_i} ,$$

$$y_{i+1} = mA_{i+1} ,$$

with

$$A_o = [1,2] , \qquad y_o = 1.5 .$$

Iterating a few times, we get

$A_1 = [1.375, 1.4375]$, $y_1 = 1.40625$,

$A_2 = [1.41406\ldots, 1.41441\ldots]$, $y_2 = 1.41424\ldots$,

$A_3 = [1.4142135\,59\ldots, 1.4142135\,66\ldots]$, $y_3 = 1.4142135\,63\ldots$.

Notice that the intervals A_0, A_1, A_2, A_3 do in fact contain $\sqrt{2} = 1.414213562\ldots$ and that the successive widths decrease like $1, 6 \cdot 10^{-2}, 4 \cdot 10^{-4}, 7 \cdot 10^{-9}$.

A given real rational function f in n variables always has extensions to rational interval functions simply by extending the real arithmetic operations in a real rational expression; but never a unique extension since, for example, we may add the expression $X - X$ to any extension and get another one with the same real restriction.

This corresponds to the fact that there are many equivalent real rational forms which are not equivalent as interval rational expressions because of the lack of distributivity and inverses in interval arithmetic.

There are, of course, extensions of a real rational function to interval valued functions on intervals which are not rational interval functions. For example, the real function defined by $f(x) = 0$, for all x, has interval extensions of the form $F([a,b]) = (b - a)[c,d]$ with domain $\mathcal{I}$.

If $[c,d] = [-1,1]$, then $F([a,b]) = [a - b, b - a] = [a,b] - [a,b]$ and F is rational; however, if $[c,d] = 1$, then $F([a,b]) = b - a$ and F is not rational.

To show that F is not a rational interval function, we can argue as follows. If F were rational, there would be a finite combination of constants and the variable interval $[a,b]$ in an expression with interval arithmetic operations representing the value of $F([a,b])$. Now F is not a constant function since, in particular, $F([0,1]) \neq F([0,2])$. But any rational interval expression in which a variable interval X appears will produce an interval of positive length when evaluated at $X = [a,b]$ for $a < b$. Since $F([a,b]) = b - a$ means that $F([a,b])$ is an interval with zero length, i.e., a real number, we can conclude there is no such rational interval expression for $F([a,b])$.

Thus, an interval valued function F may not be rational even if the left and right end points of the interval $F([a,b])$ are both rational functions of a and b. Indeed, we will consider certain such functions arising from rational interval computation in the following sections (without giving them any special name). They will arise as bounding functions for solutions to differential equations, for example. The real restriction of a rational interval function F may be an interval valued function; for example, when F is a constant function, say $F([a,b]) = [0,1]$ for all $a \leq b$.

We conclude this section with some further properties of interval arithmetic.

We have the cancellation laws:

$$(2.8) \quad \begin{cases} X + A = X + B, & \text{iff } A = B; \\ X \in \mathcal{I}^* \rightarrow AX = BX, & \text{iff } A = B. \end{cases}$$

There are many ways to see the last relation above; for the sake of variety the following argument may be used. Assume $AX = BX$; pick a number a in A and let x run through X, then for each ax, assign a pair $b(x)$, $t(x)$ such that $ax = b(x)t(x)$ with $b(x) \in B$, $t(x) \in X$. Choose $x_1 \in X$, and consider the sequence defined by $x_{i+1} = t(x_i)$, then $ax_i = b(x_i)x_{i+1}$, with $x_i \in X$.

Since X and B are compact, the x_i have a limit point in X, say $\bar{x}$, and by the continuity of real arithmetic $\bar{x}$ will satisfy

$$a\bar{x} = b(\bar{x})\bar{x} .$$

By hypothesis $\bar{x} \neq 0$, so $a = b(\bar{x}) \in B$. Similarly, $B \subseteq A$. Therefore, $A = B$. Conversely, $A = B$ obviously implies $AX = BX$.

If $A \subseteq B$, then there is an interval C such that $B = A + C$ and $0 \in C$. In fact, if $A \subseteq B$ and D is any interval, then there is an interval C such that

$$B = A + CD . \tag{2.9}$$

3. The Continuity of Interval Arithmetic.

We make $\mathcal{I}$ into a metric space with the distance function $P : \mathcal{I}^2 \to R$ (the real line) defined by $P([a,b], [c,d]) = \max (|a - c|, |b - d|)$.

Theorem 3.1.

The function P is a metric on $\mathcal{I}$.

Proof:

The positivity and symmetry of P are obvious. The triangle inequality follows easily from the definition of P, as we now show. If

$$I_1 = [a_1, b_1] , \qquad I_2 = [a_2, b_2] , \qquad I_3 = [a_3, b_3] \in \mathcal{I},$$

then

$$P(I_1, I_2) = \max (|a_1 - a_2| , |b_1 - b_2|) ,$$

$$\begin{aligned} &P(I_3, I_1) + P(I_3, I_2) \\ &= \max (|a_3 - a_1| , |b_3 - b_1|) + \max (|a_3 - a_2| , |b_3 - b_2|) \\ &= \max (|a_3 - a_1| + |a_3 - a_2| , |a_3 - a_1| + |b_3 - b_2| , \\ &\qquad |b_3 - b_1| + |a_3 - a_2| , |b_3 - b_1| + |b_3 - b_2|) . \end{aligned}$$

Now since $|a_1 - a_2| \leq |a_3 - a_1| + |a_3 - a_2|$ and $|b_1 - b_2| \leq |b_3 - b_1| + |b_3 - b_2|$, we have

$$P(I_1, I_2) \leq P(I_3, I_1) + P(I_3, I_2) .$$

Notice that we have $P([x,x], [y,y]) = |x - y|$ and the correspondence $[x,x] \to x$ is an isometry. Thus, the metric P is consistent with our identification of $[x,x]$ and x, i.e., the real line may be regarded as a subspace of the metric space $(\mathcal{I}, P)$.

If $f : I \to \mathcal{I}$ is a real valued or even an interval valued function on a closed real interval I, there is a natural extension of f to an interval valued function $\overline{f}$ on the regular domain $\mathcal{I}_I$ consisting of all the subintervals of I. In fact, $f : \mathcal{I}_I \to \mathcal{I}$ is defined by $\overline{f}(X) = \bigcup_{x \in X} f(x)$.

The value of the function $\overline{f}$ at a point X in its domain, i.e., at a subinterval X of I is the union of the real numbers or intervals in the image of X under the mapping f. If f is real valued, this is sometimes written $f(X)$. The important distinction here is that $\overline{f}$ is a single valued function. The range of $\overline{f}$ is the metric space $\mathcal{I}$ whose points are intervals of real numbers. The function $\overline{f}$ is called the <u>united extension</u> of f. (See [19], p. 552.)

More generally, if $f : I_1 \otimes I_2 \otimes \cdots \otimes I_n \to \mathcal{I}$ is an interval valued function on the Cartesian product of the closed real intervals $I_1, I_2, \ldots, I_n$, we call $\overline{f}$ the <u>united extension</u> of f (see [19]), with

$$\overline{f} : \mathcal{I}_{I_1} \otimes \mathcal{I}_{I_2} \otimes \cdots \otimes \mathcal{I}_{I_n} \to \mathcal{I} ,$$

defined by

$$\overline{f}(X_1, X_2, \ldots , X_n) = \bigcup f(x_1, \ldots , x_n) ,$$

with the union taken over

$$x_i \in X_i ; \qquad i = 1 , \ldots , n .$$

Theorem 3.2:

If $I_1, I_2, \ldots , I_n \in \mathcal{I}$ and $f : I_1 \otimes I_2 \otimes \cdots \otimes I_n \to \mathcal{I}$ is continuous, then the united extension $\overline{f} : \mathcal{I}_{I_1} \otimes \mathcal{I}_{I_2} \otimes \cdots \otimes \mathcal{I}_{I_n} \to \mathcal{I}$ is continuous.

Note: In [19] it is shown that if X, Y are compact Hausdorff spaces and $f : X \to Y$ is continuous then the united extension $\overline{f}$ is continuous. The topology used there includes ours for $\mathcal{I}$. See also [18].

Proof (of Theorem 3.2):

Let ϵ be a positive real number and let $X_i \in I_i$, $X_i' \in I_i$ for $i = 1, 2, \ldots , n$. Since f is continuous, there is a $\delta > 0$ such that $x_i \in X_i$, $x_i' \in X_i'$, $i = 1, 2, \ldots , n$, with $|x_i - x_i'| < \delta$, implies

$$\mathbf{P}\Big(f(x_1, \ldots , x_n) , f(x_1', \ldots , x_n')\Big) < \epsilon .$$

Now $P(A,B) < \epsilon$ if and only if for each $t \in A$, there is a $t' \in B$ such that $|t - t'| < \epsilon$, so

$$P\left(\overline{f}(X_1, \ldots, X_n), \overline{f}(X_1', \ldots, X_n')\right) < \epsilon ,$$

if and only if, for each

$$t \in \bigcup_{x_i \in X_i} f(x_1, \ldots, x_n) ,$$

there is a $t' \in \bigcup_{x_i' \in X_i'} f(x_1', \ldots, x_n')$ with $|t - t'| < \epsilon$. Now

$$t \in \bigcup_{x_i \in X_i} f(x_1, \ldots, x_n)$$

implies that $t \in f(x_1, \ldots, x_n)$ for some choice of $x_i \in X_i$, so $P(X_i, X_i') < \delta$ $(i = 1, 2, \ldots, n)$ implies $|x_i - x_i'| < \delta$ $(i = 1, 2, \ldots, n)$, and therefore $P(f(x_1, \ldots, x_n), f(x_1', \ldots, x_n')) < \epsilon$ for $x_i \in X_i$, $x_i' \in X_i'$. Therefore, there is a $t' \in f(x_1', \ldots, x_n')$ with $x_i' \in X_i'$ $(i = 1, 2, \ldots, n)$ such that $|t - t'| < \epsilon$, and the proof of Theorem 3.2 is complete.

Theorem 3.3:

The arithmetic operations in $\mathcal{I}$ are continuous except for division by intervals containing zero.

Proof:

The continuity of the arithmetic operations

$$+ (Y_1, Y_2) = Y_1 + Y_2 = \{y_1 + y_2 \mid y_1 \in I_1 , y_2 \in I_2\}$$

$$- (Y_1, Y_2) = Y_1 - Y_2 = \{y_1 - y_2 \mid y_1 \in I_1 , y_2 \in I_2\}$$

$$\circ (Y_1, Y_2) = Y_1 Y_2 = \{y_1 y_2 \mid y_1 \in I_1 , y_2 \in I_2\}$$

$$\div (Y_1, Y_2) = Y_1 \div Y_2 = \{y_1 / y_2 \mid y_1 \in I_1 , y_2 \in I_2 \in \mathcal{I}^*\}$$

for $I_1, I_2 \in \mathcal{I}$ follows from Th. 3.2 with $n = 2$ by the continuity of real arithmetic. Any pair of intervals I_1, I_2 is contained in the interior of $\mathcal{I}'_{I_1} \otimes \mathcal{I}'_{I_2}$ for some $I'_1, I'_2 \in \mathcal{I}$ and if $I_2 \in \mathcal{I}^*$ then I'_2 may also be chosen with $0 \notin I'_2$. Thus, except for division by intervals containing zero, the arithmetic operations are continuous. Obviously the constant functions and the "projections," i.e., functions of the type

$$F_k : \mathcal{I}^n \to \mathcal{I}$$

with

$$F_k(X_1, \ldots , X_n) = X_k$$

are continuous. We conclude the following result using Th. 3.3 and finite induction.

Theorem 3.4:

Rational interval functions are continuous.

We have already noted[*] that the extension of a real rational function to a rational interval function is never unique. The various ways of writing a rational form representing a given real rational function are equivalent in the sense of all defining the same function. For example, $x(1 - x)$, $x - x^2$, $\frac{1}{4} - (\frac{1}{2} - x)^2$, all define the same real polynomial function. On the other hand, the interval polynomial forms, $X(1 - X)$, $X - X^2$, $\frac{1}{4} - (\frac{1}{2} - X)^2$, all define different interval polynomial functions, since for $X = [0,1]$ we have

$$[0,1]\,(1 - [0,1]) = [0,1]\,[0,1] = [0,1] ,$$

$$[0,1] - [0,1]^2 = [0,1] - [0,1] = [-1, 1] ,$$

$$\frac{1}{4} - (\frac{1}{2} - [0,1])^2 = \frac{1}{4} - [-\frac{1}{2}, \frac{1}{2}]^2 = \frac{1}{4} - [-\frac{1}{4}, \frac{1}{4}] = [0, \frac{1}{2}] .$$

Furthermore, while each of these rational interval functions has as values intervals containing the range of values of the real polynomial $x(1-x)$ when $x \in X$, none of them has exactly these values. For, in fact,

$$\{x(1 - x) \mid x \in [0,1]\} = [0, \frac{1}{4}] .$$

---

*/ See pp. 8, 9, 13.

The question arises whether some rational interval function always exists which computes the exact range of values of the corresponding real rational function. We will show by a counter-example that this is not the case. We will show that there is no rational interval function which, for every interval X, has the value

$$\{x^2 \mid x \in X\} .$$

Theorem 3.5:

The function $\overline{f} : \mathcal{J} \to \mathcal{J}$ defined by $\overline{f}(X) = \{x^2 \mid x \in X\}$ is not a rational interval function.

Proof:

Let f be the polynomial function defined by $f(x) = x^2$ for all real x. Now a rational interval function F with real restriction f, $(F([x,x]) = f(x) = x^2)$ must be given by a finite interval expression in X. This expression must involve at least two occurrences of X, otherwise it is clear that x^2 will not appear as an $F([x,x])$. At the same time, if any occurrence of X in the expression is replaced by a new variable Y, the new expression will define a rational interval function $G(X,Y)$ which has the property $G(X,X) = F(X)$.

There are altogether a finite number of occurrences of X in the expression for F, so there is a rational interval function $H(X_1, X_2, \ldots, X_n)$, such that

$$H(X, X, \ldots, X) = F(X) ,$$

and such that in the expression for H, each variable X_i occurs exactly once. For such a rational interval function H it is clear that the united extension $\bar{h}$ of the real restriction gives back the function, i.e., $\bar{h} = H$.

Now consider the real restriction h of H. It is given by a real rational expression in $x_1, \dots, x_n$ with each x_i occurring exactly once and such that $h(x, x, \dots, x) = x^2$ and $\bar{h}(X_1, X_2, \dots, X_n) = H(X_1, X_2, \dots, X_n)$.

Assume that $\bar{f}(X) = \{x^2 \mid x \in X\}$ <u>is</u> a rational interval function on $\mathcal{I}$. Then there are functions h, H as described above, such that for $X_1, \dots, X_n \in \mathcal{I}$, $\bar{h}(X_1, \dots, X_n) = H(X_1, X_2, \dots, X_n)$ and for $X \in \mathcal{I}$, $H(X, \dots, X) = \bar{f}(X)$. Observe first that h must be a polynomial in $x_1, \dots, x_n$ since $H(X_1, \dots, X_n)$ is an interval for $X_1, \dots, X_n \in \mathcal{I}$.

This means that there is a real polynomial expression for $h(x_1, x_2, \dots, x_n)$ in which each x_i occurs exactly once and such that for every choice of real values for $x_1, \dots, x_n$ there is a real number x such that

$$h(x_1, x_2, \dots, x_n) = x^2 .$$

But this implies that $h(x_1, x_2, \dots, x_n) \geq 0$ for all real values of $x_1, \dots, x_n$. The expansion of h about $0, 0, \dots, 0$ in particular, must have the form $h(x, \dots, x_n) = \sum_{i \neq j} c_{ij} x_i x_j + T$ in which T contains a finite number of terms of order three and higher, and the sum after renaming variables has the form $c_1 x_1 x_2 + \cdots + c_p x_{2p-1} x_{2p}$ for some $1 \leq p \leq \frac{n}{2}$. Since $h(x, \ldots, x) = x^2$ we can see by choosing $x_i = x$ sufficiently small for $i = 1, 2, \ldots, n$ that $\sum_{k=1}^{p} c_k = 1$ and some c_k, say c_1 is different from zero. Then choose $x_i = 0$ for $i > 2$ and look at $h(x_1, x_2, 0, 0, \ldots, 0) = c_1 x_1 x_2 + T$. For $x_1, x_2 \neq 0$ but small, T is small compared to $c_1 x_1 x_2$, and $h(x_1, x_2, 0, \ldots, 0)$ takes on negative values. This contradicts $h(x_1, \ldots, x_n) \geq 0$, hence we have proved that $\bar{f}(X) = \{x^2 \mid x \in X\}$ cannot be a rational interval function.

4. Refinements.

We enlarge the class of rational interval functions by introducing a process we call refinement. The enlarged class of rational interval functions and their refinements will serve as an approximating class of functions for the united extensions of real rational functions and even for the united extensions of real restrictions of rational interval functions.

An example will serve to illustrate the simple idea involved.

The rational interval function F defined on $\mathcal{I}_{[0,1]}$ by $F(X) = X(1 - X)$ has real restriction $f : [0,1] \rightarrow$ reals, given by $f(x) = x(1 - x)$. The united extension of f is the interval function $\overline{f}$ on $\mathcal{I}_{[0,1]}$ given by $\overline{f}(X) = \{t \mid t = x(1 - x), \text{ some } x \in X\}$. We have

$$F([0,1]) = [0,1]\,(1 - [0,1]) = [0,1] ,$$

$$\overline{f}([0,1]) = [0, \tfrac{1}{4}] = \{f(x) \mid x \in X\} .$$

The united extension $\overline{f}$ of a real rational function f computes the range of values of f for arguments of f running through the argument intervals of $\overline{f}$.

Suppose we write

$$[0,1] = \bigcup_{i=1}^{n} \left[\frac{i - 1}{n}, \frac{i}{n}\right],$$

then

$$\overline{f}([0,1]) = \bigcup_{i=1}^{n} \overline{f}\left(\left[\frac{i - 1}{n}, \frac{i}{n}\right]\right) \subset \bigcup_{i=1}^{n} F\left(\left[\frac{i - 1}{n}, \frac{i}{n}\right]\right) .$$

But

$$F\left[\frac{i-1}{n}, \frac{i}{n}\right] = \left[\frac{i-1}{n}, \frac{i}{n}\right]\left(1 - \left[\frac{i-1}{n}, \frac{i}{n}\right]\right)$$

$$= \left[\left(1 - \frac{i-1}{n}\right)\left(\frac{i-1}{n}\right) - \frac{i-1}{n^2}, \left(1 - \frac{i}{n}\right)\left(\frac{i}{n}\right) + \frac{i}{n^2}\right]$$

and

$$\bigcup_{i=1}^{n} F\left(\left[\frac{i-1}{n}, \frac{i}{n}\right]\right) = [0, 1/4] + \begin{cases} \left[0, \frac{1}{2n}\right], & n \text{ even}; \\ \left[0, \frac{1}{2n} + \frac{1}{4n^2}\right], & n \text{ odd}. \end{cases}$$

We see from this, in particular, that

$$\bigcup_{i=1}^{n} F\left(\left[\frac{i-1}{n}, \frac{i}{n}\right]\right)$$

converges to $[0, 1/4]$ with increasing n.

The space $\mathcal{I}$ of intervals $[x,y]$ can be visualized as a closed half-plane in the Euclidean (x,y)-plane above and including the diagonal $y = x$. The interval $[x,y]$, which contains the real numbers t, $x \leq t \leq y$, is represented by the point with coordinates x, y. If $[u,v] \subset [x,y]$, then $[u,v]$ lies in the closed triangle $\mathcal{I}_{[x,y]}$, consisting of intervals contained in $[x,y]$ as indicated in Figure 1.

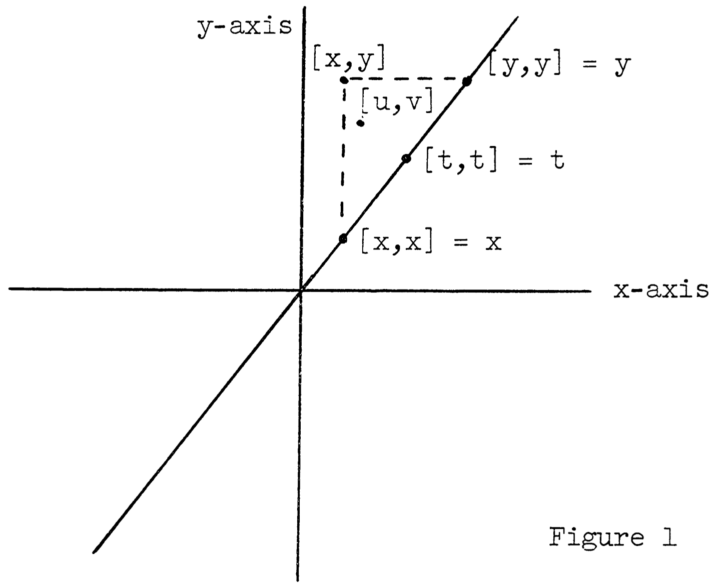

Figure 1

The width y - x of an interval [x,y] is determined by its distance from the diagonal. Intervals of constant width lie on lines parallel to the diagonal R.

Intervals whose midpoints are all a single fixed number lie on a half-line perpendicular to R.

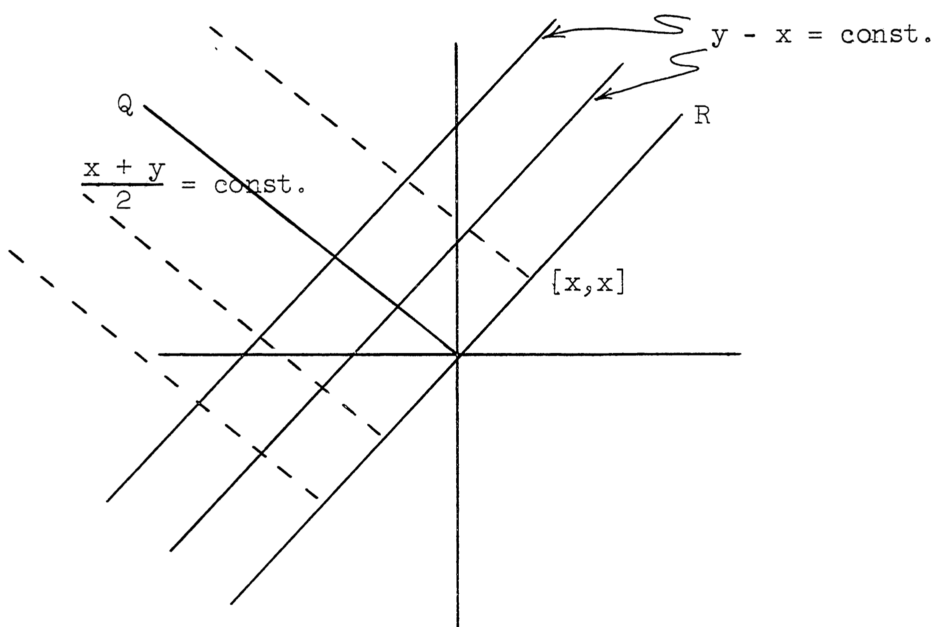

Figure 2

The negative $[-y, -x]$ of an interval $[x,y]$ is the reflection of $[x,y]$ in the half-line Q consisting of intervals with midpoint 0, (i.e., $Q = \{[-x,x] \mid x \geq 0 , x \in R\}$.

The sum of two intervals $[x_1, y_1] + [x_2, y_2] = [x_1 + x_2, y_1 + y_2]$ is constructed graphically in the same way as vectors in the plane.

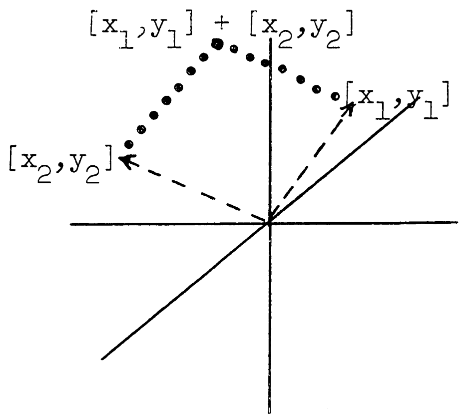

Figure 3

If $I \in \mathcal{I}$, denote by RI the union of the ray through $0, 0$ and I and its reflection in Q (i.e., the ray through $-I$).

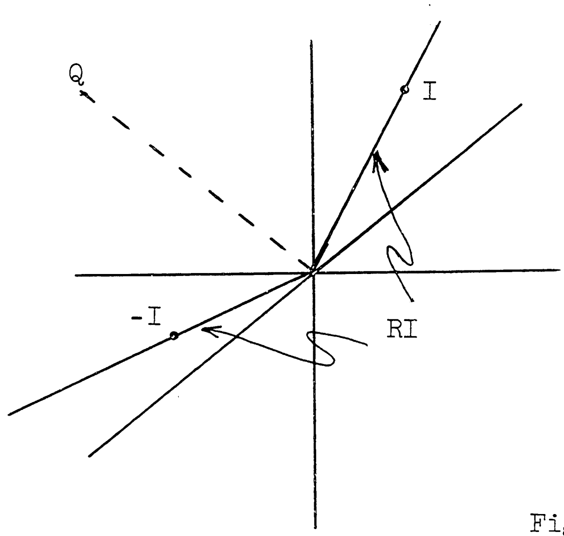

Figure 4

For $t \in R$, $I \in \mathcal{I}$, the interval tI is a segment of RI. Hence, the interval product JI, $J \in \mathcal{I}$, which can be written $JI = \bigcup_{t \in J} tI$, can be found graphically by constructing the smallest set $\mathcal{I}_A$ containing the segment of RI consisting of the points tI for $t \in J$. This turns out to be $\mathcal{I}_{JI}$.

The construction is illustrated by Figure 5, where

$$J = [-1, 3/2], \quad I = [1,2].$$

The segment of RI in question is the set of points (intervals) of the form

$$t[1,2] \quad \text{for} \quad t \in [-1, 3/2].$$

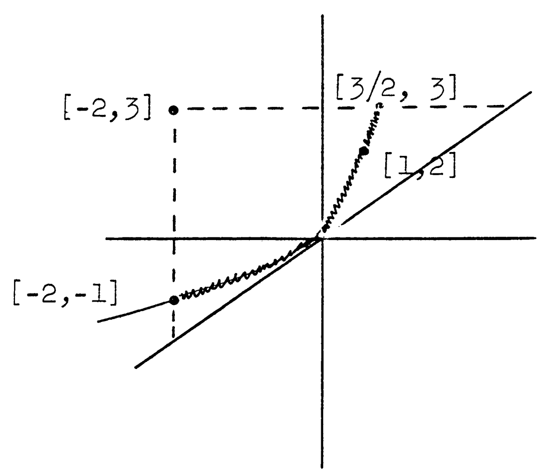

Figure 5

Suppose F is a rational interval function on $\mathcal{I}_A$. If $X \in \mathcal{I}_A$, i.e., $X \subset A$ and if

$$X = \bigcup_{i=1}^{n} X_i ,$$

with $X_i \in \mathcal{I}$, then we call

$$\bigcup_{i=1}^{n} F(X_i)$$

a <u>refinement</u> of F at X. In the geometric interpretation explained in Figure 1 the picture might look as in Figures 6 and 7.

The value of F at X is "refined" by the union of values of F at the points X_i of a finite covering of X.

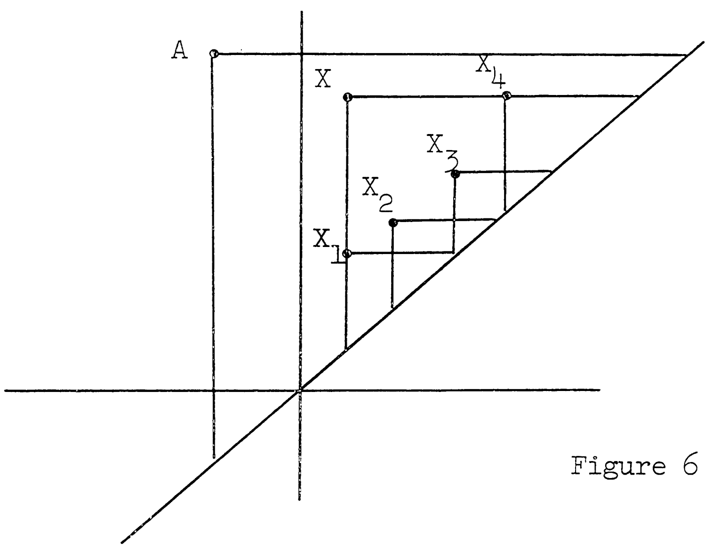

Figure 6

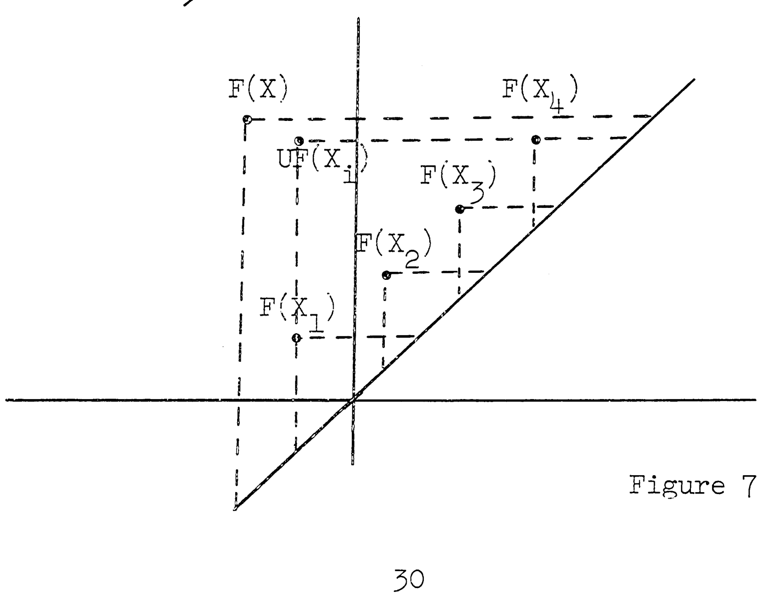

Figure 7

Since we already know that rational interval functions are continuous, and since we can make the distance $P(X_i, R)$ of X_i from the real line (i.e., the diagonal in $\mathcal{I}$) small for each $i = 1, 2, \ldots, n$ with

$$\bigcup_{i=1}^{n} X_i = X = [a,b]$$

choosing n large and say choosing

$$X_i = \left[a + (i - 1)\frac{b - a}{n} ,\ a + i\frac{(b - a)}{n} \right] ,$$

we might reasonably expect that a refinement of a rational interval function F can be made arbitrarily close to the united extension $\overline{f}$ of the real restriction f. We proceed to show that this is the case.

For convenience, we introduce the notation $w(I)$ for the width of an interval I. Thus, $w([a,b]) = b - a$. Also we write

$$|I| = \max_{x \in I} |x| .$$

So

$$|[a,b]| = \max(|a| , |b|) .$$

Clearly w is a linear functional on $\mathcal{I}$ over the positive reals; for $a, b \geq 0$, $a, b \in R$, $I, J \in \mathcal{I}$, we have

$$w(aI + bJ) = aw(I) + bw(J) .$$

We also have

$$w(IJ) = \max_{\substack{a,a' \in I \\ b,b' \in J}} |ab - a'b'|$$

and

$$|ab - a'b'| \leq |a| \, |b - b'| + |b'| \, |a - a'| ,$$

so

$$w(IJ) \leq |I| \, w(J) + |J| \, w(I) .$$

Similarly,

$$w(1/I) \leq |1/I|^2 \, w(I) \qquad (I \in \mathcal{I}^*) ,$$

$$|I + J| \leq |I| + |J| ,$$

and

$$|aI| = |a| \, |I| ,$$

$$w(aI) = |a| \, w(I) .$$

We will now show that there exists a positive real number K depending on F and A such that for every positive integer n and every

$$X = \bigcup_{i=1}^{n} X_i \subset A \quad ,$$

we have

$$\bigcup_{i=1}^{n} F(X_i) = \overline{f}(X) + E_n \; ,$$

with $0 \in E_n$ and $w(E_n) \leq K \max_i w(X_i)$ $(i = 1, \ldots , n)$.

In particular, this implies that $\overline{f}(X) \subset \bigcup_{i=1}^{n} F(X_i)$. In fact, we prove the following more general result.

Let F be a rational interval function with domain $\mathcal{I}_{A_1} \otimes \mathcal{I}_{A_2} \otimes \cdots \otimes \mathcal{I}_{A_m}$. Let f be the real restriction of F; thus, for $x_p \in A_p$, $f(x_1, \ldots , x_m) = F(x_1, \ldots , x_m)$. Let $\overline{f}$ be the united extension of f; thus, for $X_p \subset A_p$ $(p = 1, 2, \ldots , m)$, we have

$$\overline{f}(X^{(1)}, \ldots , X^{(m)}) = \{f(x_1, \ldots , x_m) \mid x_p \in X^{(p)} \quad (p = 1, 2, \ldots , m)\}.$$

Subdivide each of the intervals $X^{(p)}$ so that

$$\bigcup_{i=1}^{n} X_i^{(p)} = X^{(p)} \quad .$$

Theorem 4.1:

There is a positive real number K independent of the method of subdivision of the intervals $X^{(1)}, \ldots , X^{(m)}$, such that

$$\bigcup_{i_1, \ldots , i_m = 1}^{n} F\,(X_{i_1}^{(1)}, \ldots , X_{i_m}^{(m)}) = \overline{f}\,(X^{(1)}, \ldots , X^{(m)}) + E_n \; ,$$

with $0 \in E_n$ and $w(E_n) \leq K \max_{p,i} w(X_i^{(p)})$.

Proof:

We need only show the convergence part of the statement; the rest is merely restatement of previous discussion. That is, we must show that

$$w(E_n) \leq K \max_{p,i} w(X_i^{(P)}) .$$

It is clear that for the united extension $\overline{f}$ we have

$$\overline{f}(X^{(1)}, \ldots , X^{(m)}) = \bigcup_{i_1,\ldots,i_m=1}^{n} \overline{f}(X_{i_1}^{(1)}, \ldots , X_{i_m}^{(m)})$$

and we already know that there is an $E_{i_1,\ldots,i_m}$ such that

$$F(X_{i_1}^{(1)}, \ldots , X_{i_m}^{(m)}) = \overline{f}(X_{i_1}^{(1)}, \ldots , X_{i_m}^{(m)}) + E_{i_1,\ldots,i_m} ,$$

with $0 \in E_{i_1,\ldots,i_m}$. If $E_n = \bigcup E_{i_1,\ldots,i_m}$, then $|E_n| = \max|E_{i_1,\ldots,i_m}|$ and $w(E_n) \leq 2|E_n|$, so it is sufficient to show that

$$(4.1) \quad \begin{aligned} |E_{i_1,\ldots,i_m}| &= P\left(F(X_{i_1}^{(1)}, \ldots , X_{i_m}^{(m)}) , \overline{f}(X_{i_1}^{(1)}, \ldots , X_{i_m}^{(m)})\right) \\ &\leq K \max_{p,i} w(X_i^{(p)}) . \end{aligned}$$

In the expression for $F(X^{(1)}, \ldots , X^{(m)})$ each variable $X^{(p)}$ occurs only a finite number of times (possibly zero). In each occurrence substitute a new variable $X_j^{(p)}$, $j = 1, 2, \ldots , J_p$.

After substitution of the variables $x_j^{(p)}$ we obtain a new expression which we will call $H(X_1^{(1)}, \dots, X_{J_1}^{(1)}, X_1^{(2)}, \dots, X_{J_m}^{(m)})$. There are also a finite number of interval constants $C_1, \dots, C_q$ in the expression for $F(X^{(1)}, \dots, X^{(m)})$.

For each choice of real numbers $c_1, \dots, c_q$ from the intervals $C_1, \dots, C_q$ with $c_i \in C_i$, there is a real rational function which we will denote by h_c for $c = (c_1, \dots, c_q)$ and whose values satisfy $\bigcup_c h_c(x_1^{(1)}, \dots, x_{J_m}^{(1)}) = \overline{h}(x_1^{(1)}, \dots, x_{J_m}^{(m)})$ where $\overline{h}$ is the united extension of h, the real restriction of H, with the union taken over all $c = (c_1, \dots, c_q)$ with $c_i \in C_i$, $i = 1, \dots, q$. The set of rational functions h_c is uniformly bounded (i.e., for $c_i \in C_i$) and has a uniform Lipschitz constant, i.e., there is a real number K independent of c such that for all c with $c_i \in C_i$, $i = 1, \dots, q$, and for all $x^{(p)}$, $x_j^{(p)} \in X^{(p)}$, $p = 1, \dots, m$, $j = 1, \dots, J_p$,

$$
\begin{aligned}
&\Big| \, h_c\left(x_1^{(1)},\dots, x_{J_1}^{(1)},\dots, x_1^{(m)},\dots, x_{J_m}^{(m)}\right) \\
(4.2) \qquad & - h_c\left(x^{(1)},\dots, x^{(1)},\dots, x^{(m)},\dots, x^{(m)}\right) \Big| \le K \max_{j,p} |x_j^{(p)} - x^{(p)}| \, .
\end{aligned}
$$

Now

$$
\begin{aligned}
F(X^{(1)},\dots, X^{(m)}) &= H(X^{(1)},\dots, X^{(1)},\dots, X^{(m)},\dots, X^{(m)}) \\
&= \bigcup_{x_j^{(p)} \in X^{(p)}} \overline{h}(x_1^{(1)},\dots, x_{J_1}^{(1)},\dots, x_1^{(m)},\dots, x_{J_m}^{(m)}) \\
&\qquad\qquad p = 1, \dots, m \, ; \quad j = 1, \dots, J_p
\end{aligned}
$$

$$\bar{f}(X^{(1)}, \ldots, X^{(m)}) = \bar{h}(X^{(1)}, \ldots, X^{(1)}, \ldots, X^{(m)}, \ldots, X^{(m)})$$

and

$$\sup_{x_j^{(p)}, x^{(p)} \in X^{(p)}} |x_j^{(p)} - x^{(p)}| = w(X^{(p)}) .$$

Therefore, substituting $X_{i_p}^{(p)}$ for $X^{(p)}$ in (4.2), we get (4.1), so

$$P\left(F(X^{(1)}, \ldots, X^{(m)}), \bar{f}(X^{(1)}, \ldots, X^{(m)})\right) \leq K \max_{p,i} w(X_i^{(p)}) .$$

This completes the proof of Theorem 4.1.

In the special case of $m = 1, n = 1,$ and for real valued real restriction f, we have

$$F(X) = \bar{f}(X) + E$$

with

$$w(E) \leq K\, w(X)$$

and, since f satisfies a Lipschitz condition in A, we have the result:

<u>Theorem</u> 4.2:

For any rational interval function with (regular) domain $\mathcal{I}_A$ and real valued real restriction there exists a real number K such that $X \subset A$ implies $w(F(X)) \leq K\, w(X)$.

5. <u>Interval Integrals</u>.

If f is a real rational function, then the indefinite integral $\int_a^y f(x)dx$ can be expressed in terms of elementary functions of y by factoring the denominator of f and using the partial fraction decomposition of f. For an actual numerical evaluation this requires, of course, the determination of the roots of the denominator polynomial of f.

We consider in this section a more direct approach using interval computations.

Suppose $F : \mathcal{I}_A \to \mathcal{I}$ is a rational interval function with real-valued real restriction $f = F|A$. Then f is a bounded real rational function on the interval A.

If $X = [a,b] \subset A$, then $f(x) \in F(X)$ for all $x \in X$, and

$$\int_a^b f(x)dx = f(a + \theta(b - a))(b - a) ,$$

for some $\theta \in [0,1]$. Therefore,

$$\int_a^b f(x)dx \in F(X)(b - a) .$$

Now suppose $Y = [a,y] \subset A$; define

$$Y_i^{(n)} = a + [i - 1, i] \frac{(y - a)}{n} ,$$

then $Y_i^{(n)} \subset A$, and

$$Y = \bigcup_{i=1}^{n} Y_i^{(n)} .$$

By the additivity of the integral, we have

$$\int_a^y f(x)dx \in \sum_{i=1}^{n} F(Y_i^{(n)}) \frac{(y - a)}{n} \qquad \text{[interval sum]} .$$

More generally, if $Y_i^{(n)}$, $i = 1, 2, \dots, n$, is a collection of intervals such that

$$Y_i^{(n)} \subset A ,$$

$$Y = \bigcup_{i=1}^{n} Y_i^{(n)} ,$$

and

$$w(Y) = \sum_{i=1}^{n} w(Y_i^{(n)}) ,$$

then

$$\int_a^y f(x)dx \in \sum_{i=1}^{n} F(Y_i^{(n)}) \, w(Y_i^{(n)}) .$$

The intervals $Y_i^{(n)}$ do not all have to have the same width. From Theorem 4.2, it follows that there exists a K such that for $X \subset A$, we have

$$w\left(\sum_i F(Y_i^{(n)}) \, w(Y_i^{(n)})\right) \leq (y - a) \, K \max_i w(Y_i^{(n)}) ,$$

$$i = 1 , \dots , n ,$$

and we have proved that

Theorem 5.1:

$$\sum_{i=1}^{n} F(Y_i^{(n)})\ w(Y_i^{(n)}) = \int_a^y f(x)dx + E \tag{5.1}$$

with $0 \in E$, and

$$w(E) \leq (y - a)\ K \max w(Y_i^{(n)}) .$$

For an illustrative example, consider

$$\ell n\ y = \int_1^y \frac{dx}{x} , \qquad y > 1 ,$$

and take $F(Y) = 1/Y$ for $Y \geq 1$ (i.e., $y \in Y \rightarrow y \geq 1$). Let

$$Y_i^{(n)} = 1 + [i - 1,\ i]\ \frac{(y - 1)}{n} , \qquad i = 1 , 2 , \ldots , n ,$$

then

$$f(Y_i^{(n)}) = \frac{1}{1 + [i - 1,\ i]\ \frac{(y - 1)}{n}} = \left[\frac{n}{n + i(y - 1)} , \frac{n}{n + (i - 1)(y - 1)} \right] .$$

If $a \geq 1$, then

$$w(F([a,b])) = w[1/b,\ 1/a] = \frac{b - a}{ab} \leq w[a,b] ,$$

so we can take $K = 1$ in (5.1) and obtain for $n = 1, 2, \dots,$

$$\ln y \in I_n(y) \quad \text{and} \quad w(I_n(Y)) \leq \frac{y-1}{n} .$$

From (5.1), it follows that

$$(5.2) \qquad \bigcap_{n=1}^{\infty} \sum_{i=1}^{n} F(Y_i^{(n)})\, w(Y_i^{(n)}) = \int_a^y f(x)dx .$$

We have so far in this section considered only rational interval functions F whose real restrictions are real valued. If we drop this requirement we can still form the sums

$$\sum_{i=1}^{n} f(Y_i^{(n)})\, w(Y_i^{(n)}) ,$$

and we can regard (5.2) as a definition of the integral on the right hand side. Again, using Theorem 4.2, we can prove (5.1) for interval valued functions F satisfying

$$X' \subset X \rightarrow F(X') \subset F(X)$$

and

$$w(F(X)) < K\, w(X)$$

in a regular domain (see [13], [14]).

The difference between this and the previous interpretation of (5.1) should be made clear by some examples.

Let F be the rational interval function given by $F(X) = (A + X)A$ for some constant interval A and say for $X \subset [1,2]$. Subdivide the interval $[1,2]$ so that

$$Y_i^{(n)} = 1 + [i - 1, i]\frac{1}{n} \quad \text{for} \quad i = 1, 2, \ldots, n,$$

and consider the interval sums

$$I_n = \sum_{i=1}^{n} F(Y_i^{(n)})\, w(Y_i^{(n)})$$

$$= \sum_{i=1}^{n} \frac{1}{n}(A + 1 + [i - 1, i]\frac{1}{n})A .$$

Now the real restriction of F is $f = F \mid [1,2]$; so $f(x) = F([x,x]) = (A + x)A$ for $x \in [1,2]$, according to (5.2)

$$\int_1^2 (A + x)A\, dx = \bigcap_{n=1}^{\infty} \sum_{i=1}^{n} \frac{1}{n}(A + 1 + [i - 1, i]\frac{1}{n})A .$$

It is clear that the set of numbers

$$\int_1^2 (a + x)a\, dx = a(a + 3/2)$$

for $a \in A$, is contained in the interval

$$\int_1^2 (A + x)A\, dx ;$$

however, if $A = [-1, 0]$, then $\{a(a + 3/2) \mid a \in [-1, 0]\} = [-9/16, 0]$, whereas

$$\sum_{i=1}^{n} \frac{1}{n}\left([-1, 0] + 1 + [i - 1, i]\frac{1}{n}\right) [-1, 0]$$

$$= \sum_{i=1}^{n} \frac{1}{n}\left(\left[\frac{i-1}{n}, 1 + \frac{i}{n}\right]\right) [-1, 0]$$

$$= \sum_{i=1}^{n} \frac{1}{n}\left[-1 - \frac{i}{n}, 0\right]$$

$$= -\frac{1}{n^2}\left[0, \sum_{i=1}^{n} (n + i)\right] = -\frac{1}{n^2}\left[0, n^2 + \frac{n(n+1)}{2}\right]$$

$$= \left[-\frac{3}{2} - \frac{1}{2n}, 0\right]$$

and

$$\int_1^2 ([-1, 0] + x) [-1, 0]dx = \bigcap_{n=1}^{\infty}\left[-\frac{3}{2} - \frac{1}{2n}, 0\right] = \left[-\frac{3}{2}, 0\right].$$

Of course, by refining the constant A, we can approach the correct set of values for

$$\left\{ \int_1^2 (a + x)a \, dx \mid a \in A \right\}$$

that is, for

$$A = \bigcup_{j=1}^{m} A_j$$

we will obtain

$$\bigcup_{j=1}^{m} \int_1^2 (A_j + x)A_j \, dx \to \left\{\int_1^2 (a + x)a \, dx \mid a \in A\right\},$$

for

$$\max_j w(A_j) \rightarrow 0 , \qquad j = 1 , \ldots , m .$$

In the example following (5.1), we had

$$\ell n\ y \in I_n(y) = \sum_{i=1}^{n} F(Y_i^{(n)}(y))\ w(Y_i^{(n)}(y)) ,$$

with

$$F(Y) = \frac{1}{Y}$$

and

$$Y_i^{(n)}(y) = 1 + [i - 1, i] \frac{y - 1}{n} ,$$

where we found that

$$w(I_n(Y)) \leq \frac{y - 1}{n} .$$

Recall also that

$$w(F(Y)) \leq w(Y) \quad \text{for} \quad Y \geq 1 .$$

The interval valued functions,

$$Y_i^{(n)}(y) \ , \quad I_n(y) \ ,$$

of the real variable y can be extended to rational interval functions on $Y' \geq 1$ in the obvious way

$$Y_i^{(n)}(Y') \ = \ 1 + [i - 1, i] \frac{Y' - 1}{n} \ ,$$

$$I_n(Y') \ = \ \sum_{i=1}^{n} F(Y_i^{(n)}(Y')) \ w(Y_i^{(n)}(Y')) \ .$$

Clearly, $\{\ell n\ y \mid y \in Y'\} \subset I_n(Y')$ and

$$\int_{Y'} \ell n\ y\ dy \ \in \ I_n(Y') \ w(Y') \ .$$

Suppose $Y' = [y, y']$ with $y \geq 1$. Let

$$Y_j^{(m)} \ = \ y + [j - 1, j] \frac{(y' - y)}{m} \ ,$$

then

$$Y' \ = \ \bigcup_{j=1}^{m} Y_j^{(m)}$$

and

$$\sum_{j=1}^{m} w(Y_j^{(m)}) \ = \ w(Y') \ .$$

Now $I_n(Y')$ is a rational interval function of Y', so by Theorem 4.1 there exists a positive real number K such that

$$\bigcup_{j=1}^{m} I_n(Y_j^{(m)}) = \overline{I}_n(Y') + E_m$$

with $0 \in E_m$ and $w(E_m) \leq K \frac{(y' - y)}{m}$. Since $w(I_n(y)) \leq \frac{y - 1}{n}$, and $\ell n\ y \in I_n(y)$, we have

$$\overline{I}_n(Y') = \bigcup_{y \in Y'} I_n(y) \subset \{\ell n\ y \mid y \in Y'\} + [-1, 1] \frac{y' - 1}{n} .$$

Therefore

$$\bigcup_{j=1}^{m} I_n(Y_j^{(m)}) = \{\ell n\ y \mid y \in Y'\} + E_{n,m} ,$$

with $0 \in E_{n,m}$, $w(E_{n,m}) \leq \frac{K\, w(Y')}{m} + \frac{2(y' - 1)}{n}$. It is also easily shown that there are positive real numbers K, K', such that

$$\sum_{j=1}^{m} I_n(Y_j^{(m)})\ w(Y_j^{(m)}) = \int_{Y'} \ell n\ y\ dy + E'_{n,m} ,$$

with $0 \in E'_{n,m}$ and

$$w(E'_{n,m}) \leq \frac{K\, w(Y') + K'\{w(Y')\}^2}{m} + \frac{2\, w(Y')(y' - 1)}{n} .$$

Thus by iterating and composing the processes of refinement and interval integration we can obtain, by finite computations with intervals, sequences of intervals containing and converging to the range of values and the integrals of real valued functions such as the logarithm which are not themselves rational but which are integrals of rational functions.

We consider now some more rapidly convergent procedures for bounding real integrals with interval computations.

Suppose $F^{(0)}, F^{(1)}, \ldots, F^{(k)}$ are rational interval functions on $\mathcal{I}_A$ such that the corresponding real restrictions are real valued rational functions $f^{(r)} = F^{(r)} \mid A$ with $f^{(r)}(x) = \frac{d}{dx} f^{(r-1)}(x)$, $r = 1, \ldots, k$. That is, the $\{f^{(r)}\}$ are the successive derivatives up to order k of an ordinary rational function with real coefficients, namely $F^{(0)} \mid A = f^{(0)} = f$. Furthermore, f is bounded on A. In fact, $x \in A$ implies $f(x) \in F^{(0)}(A)$. Consider the real integral

$$\int_a^b f(x)\, dx \quad ,$$

with $[a,b] \subset A$. Subdivide the interval $[a,b]$ so that

$$[a,b] = \bigcup_{i=1}^{n} X_i \; ,$$

with

$$\sum_{i=1}^{n} w(X_i) = b - a \; ,$$

and write $X_i = [x_{i-1}, x_i]$ and $f = f^{(0)}$. The Taylor theorem with remainder asserts that for each $t \in [0, w(X_i)]$,

$$f(x_{i-1} + t) = f^{(0)}(x_{i-1}) + f^{(1)}(x_{i-1})t + \cdots + \frac{f^{(k-1)}(x_{i-1})}{(k-1)!} t^{k-1} + R_{i-1}^{(k)}(t)$$

with

$$R_{i-1}^{(k)}(t) = \frac{1}{k!} f^{(k)}(x_{i-1} + \theta_t\ t)\ t^k$$

for some $\theta_t \in [0,1]$. Now

$$\int_{X_i} f(x)\ dx = \sum_{r=0}^{k-1} \frac{f^{(r)}(x_{i-1})}{r!} \int_0^{w(X_i)} t^r\ dt + \int_0^{w(X_i)} R_{i-1}^{(k)}(t)\ dt$$

and the last integral exists since all the others do. We can write

$$\int_0^h g(t)\ t^k\ dt = \frac{1}{k+1} \int_0^h g(t)\ d(t^{k+1})\ ;$$

therefore,

$$\int_0^{w(X_i)} R_{i-1}^{(k)}(t)\ dt$$

$$= \frac{1}{k!} \int_0^{w(X_i)} f^{(k)}(X_{i-1} + \theta_t\ t)t^k\ dt \in \frac{1}{(k+1)!} F^{(k)}(X_i)\ \{w(X_i)\}^{k+1}\ .$$

Since

$$\int_0^{w(X_i)} t^r\ dt = \frac{1}{r+1} \{w(X_i)\}^{r+1}\ ,$$

we have finally the result that

$$\int_a^b f(x)dx \in \sum_{i=1}^{n} \sum_{r=0}^{k-1} \frac{f^{(r)}(x_{i-1})}{(r+1)!} \{w(X_i)\}^{r+1} + E_{n,k}$$

with

$$E_{n,k} = \frac{1}{(k+1)!} \sum_{i=1}^{n} F^{(k)}(X_i)\ \{w(X_i)\}^{k+1}\ .$$

Since there is a K_k such that for all $X_i \subset A$,

$$w\left(F^{(k)}(X_i)\right) \leq K_k \, w(X_i) ,$$

it follows that

$$w(E_{n,k}) \leq \frac{K_k}{(k+1)!} (b - a) \max_i \{w(X_i)\}^{k+1} ,$$

$$i = 1 , \ldots , n .$$

Now define $I_{n,k}$, $n,k \geq 1$, by

$$(5.3) \quad I_{n,k} = \sum_{i=1}^{n} \sum_{r=0}^{k-1} \frac{F^{(r)}(x_{i-1})}{(r+1)!} \{w(X_i)\}^{r+1} + \frac{1}{(k+1)!} \sum_{i=1}^{n} F^{(k)}(X_i)\{w(X_i)\}^{k+1} .$$

We have proved that

<u>Theorem</u> 5.4:

$$\int_a^b f(x)\,dx \in I_{n,k} , \qquad n , k \geq 1 ,$$

and if $w(X_i) \leq h$ for $i = 1 , \ldots , n$, then

$$(5.4) \qquad w(I_{n,k}) \leq \frac{K_k}{(k+1)!} (b - a) \, h^{k+1} .$$

The formula (5.3) gives a $(k+1)^{st}$ order method in the sense of (5.4) for each positive integer k. In case $k = 0$, delete the double sum on the right hand side of (5.3) and the first order method expressed by (5.1) results.

For an example, consider

$$\int_1^2 \frac{dx}{x} .$$

Let $X_i = 1 + [i - 1, i] \frac{1}{n}$, $i = 1, \ldots, n$. We have $f^{(0)}(x) = f(x) = \frac{1}{x}$, so

$$f^{(r)}(x) = \frac{(-1)^r \, r!}{x^{r+1}} , \qquad r = 0, 1, 2, \ldots .$$

Now take

$$F^{(r)}(X) = \frac{(-1)^r \, r!}{X^{r+1}} , \qquad r = 0, 1, 2, \ldots ,$$

then

$$w\left(F^{(r)}(X)\right) = r! \, w\left(\frac{1}{X^{r+1}}\right) .$$

If $X = [a,b] \subset [1,2]$, then

$$\frac{1}{[a,b]^{r+1}} = \left[\frac{1}{b^{r+1}} , \frac{1}{a^{r+1}}\right]$$

and

$$w\left(\frac{1}{X^{r+1}}\right) = \frac{1}{a^{r+1}} - \frac{1}{b^{r+1}} = (b - a) \frac{(b^r + \cdots + a^r)}{a^{r+1} \, b^{r+1}}$$

$$\leq \frac{(r + 1)}{ba^{r+1}} \, w(X) \leq (r + 1) \, w(X) .$$

Thus we can use $K_k = (k + 1)!$, $b - a = 1$, $h = \frac{1}{n}$ in (5.4) to obtain

$$\int_1^2 \frac{dx}{x} \in I_{n,k} ,$$

with

$$w(I_{n,k}) \leq (\tfrac{1}{n})^k , \tag{5.5}$$

where

$$\begin{aligned} I_{n,k} = & \sum_{i=1}^{n} \sum_{r=0}^{k-1} \frac{(-1)^r}{r + 1} \left(1 + \frac{i - 1}{n}\right)^{-r-1} (\tfrac{1}{n})^{r+1} \\ & + \frac{1}{k + 1} \sum_{i=1}^{n} (-1)^k \left(1 + [i - 1, i] \frac{1}{n}\right)^{-k-1} (\tfrac{1}{n})^{k+1} . \end{aligned} \tag{5.6}$$

Call $y_i = \frac{1}{n + i - 1}$; then (5.6) can be rewritten as

$$\begin{aligned} I_{n,k} = & \sum_{i=1}^{n} \left\{ y_i \left(1 + y_i \left(- \frac{1}{2} + \cdots + y_i \left(\frac{(-1)^{k-1}}{k} \right) \cdots \right) \right) \right\} \\ & + \frac{(-1)^k}{k + 1} \sum_{i=1}^{n} [n + i - 1, n + i]^{-k-1} . \end{aligned} \tag{5.7}$$

We now give a heuristic discussion of the "efficiency" of this formula. In the final section after rounded interval arithmetic has been introduced we will reconsider briefly the problem of efficiency from the point of view of actual machine computation.

Using (5.7), the computation of $I_{n,k}$ requires roughly $3kn$ additions and multiplications of real numbers (ignoring about $n + 2$

divisions). Looking at the bound (5.5), suppose we wish to make

$$(\frac{1}{n})^k = \epsilon$$

then

$$k = \frac{\ell n \frac{1}{\epsilon}}{\ell n\ n} .$$

The quantity $C_{n,k} = 3kn$ measures the amount of computation required to evaluate $I_{n,k}$. In order to achieve $w(I_{n,k}) \leq \epsilon$ it is sufficient (and in order to guarantee it, it is necessary) to use any positive integer n together with k_n, the smallest integer k satisfying

$$k \geq \frac{\ell n \frac{1}{\epsilon}}{\ell n\ n} .$$

Then the amount of computation required will be $C_{n,k_n} = 3k_n n$ or very nearly

$$C(n) = \frac{3n}{\ell n\ n} \ell n \frac{1}{\epsilon} .$$

The function $C(n)$ has a minimum at $n = 3$ for positive integers n, so the most efficient choice of k, n indicated by this argument is $n = 3$, and k the smallest integer satisfying

$$k \geq \frac{\ell n \frac{1}{\epsilon}}{\ell n\ 3}$$

in which case we find that the amount of computation required for

$$(\frac{1}{n})^k = \epsilon ,$$

is very nearly

$$C(3) = (8.19 \ldots) \ln \frac{1}{\epsilon} .$$

If $\epsilon = 10^{-10}$ for example, we choose $n = 3$, $k = 21 > \frac{\ln 10^{10}}{\ln 3} = 20.9 \ldots$ and using (5.7) to compute $I_{3,21}$ we would have $C_{3,21} = 189$, and

$$w(I_{3,21}) = w\left(\frac{1}{22} \sum_{i=1}^{3} [3 + i - 1, 3 + i]^{-22}\right)$$

$$= \frac{1}{22} \left\{ w([3,4]^{-22}) + w([4,5]^{-22}) + w([5,6]^{-22}) \right\}$$

$$= \frac{1}{22} \left\{ (\frac{1}{3})^{22} - (\frac{1}{6})^{22} \right\} < 10^{-10}$$

as claimed. Suppose we arbitrarily choose a value of k, say $k = 4$. Then, to guarantee $w(I_{n,4}) \leq 10^{-10}$, with this 4^{th} order method we need to take n according to (5.5) such that

$$(\frac{1}{n})^4 \leq 10^{-10} ,$$

or $n \geq 317$ and in this case $C_{317,4} = 3 \cdot 317 \cdot 4 = 3804$. Or, in other words, this choice requires about 20 times as much computation as our previous choice, $n = 3$, $k = 21$.

The method defined by (5.3) was based on a local expansion of the integrand $f(x)$ in Taylor's series and was introduced mainly as a straightforward illustration of the derivation of "high order convergent" procedures for bounding definite integrals with interval arithmetic.

There are, of course, many other rational expressions approximating

$$\int_a^b f(x)\,dx$$

such that the error can be bounded by rational interval computations.

A highly efficient procedure in widespread use is the so-called Gaussian quadrature technique [11]. We write

$$[a,b] = \bigcup_{i=1}^{n} X_i$$

with

$$\sum_{i=1}^{n} w(X_i) = b - a$$

with the same assumptions on f as in (5.3). Then $X_i = [x_{i-1}, x_i]$ and the Gaussian method has the form

$$(5.8) \qquad \int_a^b f(x)dx = \sum_{i=1}^{n} \left\{ w(X_i) \sum_{r=1}^{k} g_r\, f\Big(x_{i-1} + u_r(w(X_i))\Big) \right\} + E_{n,k}$$

where

$$(5.9) \qquad E_{n,k} = \frac{(k!)^4}{[(2k)!]^3\,(2k+1)} \sum_{i=1}^{n} \{w(X_i)\}^{2k+1} f^{(2k)}(\xi_i)$$

for some $\xi_i \in X_i \quad i = 1, 2, \ldots, n.$

The numbers g_r and u_r $(r = 1, 2, \ldots, k)$ are the weights and argument spacings of the Gauss "k-point formula" [11]. They are associated with the zeros of the Legendre polynomials

$$P_k(t) = \frac{d^k}{dt^k}(t^2 - 1)^k$$

and are tabulated to 15 decimal place accuracy for $k = 1, \ldots, 16$ in [10].

Using Stirling's inequalities for $n!$ we find that for positive integer values of k:

$$\frac{2k}{2k+1} 2\pi \left(\frac{1}{4}\right)^{2k+1} < \frac{[k!]^4}{[(2k)!]^2 (2k+1)} < 2\pi \left(\frac{1}{4}\right)^{2k+1} .$$

Since $f^{(2k)}(\xi_i) \in F^{(2k)}(X_i)$ for a rational interval function $F^{(2k)}$ with real restriction $f^{(2k)}$, we can write (using (5.9))

$$E_{n,k} \in 2\pi \left[\frac{2k}{2k+1}, 1\right] \sum_{i=1}^{n} \left(\frac{w(X_i)}{4}\right)^{2k+1} \frac{F^{(2k)}(X_i)}{(2k)!} .$$

We now define an interval version of the Gaussian method by the formula

$$\text{(5.10)} \quad \begin{aligned} I^G_{n,k} = &\sum_{i=1}^{n} w(X_i) \sum_{r=1}^{k} g_r\, f(x_{i-1} + u_r\, w(X_i)) \\ &+ 2\pi \left[\frac{2k}{2k+1}, 1\right] \sum_{i=1}^{n} \left(\frac{w(X_i)}{4}\right)^{2k+1} \frac{F^{(2k)}(X_i)}{(2k)!} . \end{aligned}$$

For this method we have

$$\int_a^b f(x)\, dx \in I^G_{n,k} ;$$

for $X \subset A$, there is a K_{2k} such that $w(F^{(2k)}(X)) \leq K_{2k}\, w(X)$, and putting $h = \max_i w(X_i)$, $i = 1, 2, \ldots, n$, we have

$$(5.11) \quad w(I^G_{n,k}) \leq \frac{2\pi}{(2k)!}(b - a)\left\{\frac{h}{4} K_{2k} + \frac{1}{4(2k + 1)} \left|F^{(2k)}([a,b])\right|\right\} \left(\frac{h}{4}\right)^{2k} .$$

<u>Proof</u> of (5.11):

Recall that $w(AB) \leq |A|\, w(B) + |B|\, w(A)$ and $w(aA + bB) = |a|\, w(A) + |b|\, w(B)$. Thus,

$$w(I^G_{n,k}) \leq 2\pi \left\{ \left(\frac{h}{4}\right)^{2k+1} \frac{K_{2k}(b - a)}{(2k)!} + \frac{1}{(2k + 1)(2k)!} \sum_{i=1}^{n} \left(\frac{w(X_i)}{4}\right)^{2k+1} \left|F^{(2k)}(X_i)\right| \right\} .$$

Now

$$\sum_{i=1}^{n} \left(\frac{w(X_i)}{4}\right)^{2k+1} \left|F^{(2k)}(X_i)\right| \leq \frac{1}{4}\left(\frac{h}{4}\right)^{2k} \sum_{i=1}^{n} \left|F^{(2k)}(X_i)\right| w(X_i)$$

and

$$\sum_{i=1}^{n} \left|F^{(2k)}(X_i)\right| w(X_i) \leq \left|F^{(2k)}([a,b])\right| (b - a) .$$

Putting these inequalities together we obtain (5.11).

Returning to our example

$$\int_1^2 \frac{dx}{x} ,$$

put

$$F^{(2k)}(X) = \frac{(-1)^{2k}\,(2k)!}{X^{2k+1}}$$

as before. We can take $K_{2k} = (2k+1)!$ and by direct computation we find that

$$|F^{(2k)}([1,2])| = (2k)! \quad .$$

So with $w(X_i) = \frac{1}{n}$ we have

$$\int_1^2 \frac{dx}{x} \in I^G_{n,k}$$

with (5.10) becoming

$$(5.12) \qquad \begin{aligned} I^G_{n,k} &= \sum_{i=1}^{n} \frac{1}{n} \sum_{r=1}^{k} g_r \, \frac{1}{1 + \frac{i-1}{n} + u_r \frac{1}{n}} \\ &+ 2\pi \left[\frac{2k}{2k+1} , 1 \right] \sum_{i=1}^{n} \left(\frac{1}{4n}\right)^{2k+1} \frac{(-1)^{2k}}{\left[1 + \frac{i-1}{n}, 1 + \frac{i}{n}\right]^{2k+1}} \end{aligned}$$

and (using (5.11))

$$(5.13) \qquad w(I^G_{n,k}) \leq 2\pi \left\{ \frac{1}{4n}(2k+1) + \frac{1}{4(2k+1)} \right\} \left(\frac{1}{4n}\right)^{2k} \quad .$$

Counting a division as 3 multiplications, the number of multiplications and additions to evaluate $I^G_{n,k}$ by (5.12) is roughly (ignoring n additions)

$$C^G_{n,k} = (4k+8)\,n \quad .$$

In order to achieve $w(I_{n,k}) \leq \epsilon$ it is sufficient to take n, k positive integers such that

$$2\pi \left\{ \frac{1}{4\pi} (2k + 1) + \frac{1}{4(2k + 1)} \right\} \left(\frac{1}{4n}\right)^{2k} \leq \epsilon .$$

If $\epsilon = 10^{-10}$, we can choose $n = 1$ and $k = 10$, in which case $C^G_{1,10} = 48$. This is evidently the most efficient choice of n, k in this example; if $w(I^G_{n,k}) \leq 10^{-10}$ and $C^G_{n,k} \leq 48$, then $n = 1$, $k = 10$. Recall that our best choice of n, k for $I_{n,k}$ with the same value for ϵ was $n = 3$, $k = 21$, in which case $C_{3,21}$ was 189. In other words, it takes about a fourth as many arithmetic operations to evaluate

$$\int_1^2 \frac{dx}{x}$$

using $I^G_{n,k}$ as it does using $I_{n,k}$ to achieve guaranteed ten decimal place accuracy.

6. The Initial Value Problem in Ordinary Differential Equations.

In this section we are interested in the problem of computing bounds on the solution to the system of first order ordinary differential equations,

$$(6.1) \qquad \frac{dy_j}{dx} = f_j(x, y_1, \ldots, y_m), \qquad j = 1, \ldots, m,$$

which satisfies the initial conditions

$$(6.2) \qquad y_j(x_o) = y_{jo}, \qquad j = 1, \ldots, m.$$

For brevity, we will sometimes use the vector notation y for $(y_1, \ldots, y_m)$ and f for $(f_1, \ldots, f_m)$. For example, we can write (6.1) in the simpler form

$$(6.3) \qquad \frac{dy}{dx} = f(x,y),$$

and (6.2) can be written $y(x_o) = y_o$. We will use the metric $|y - z| = \max\{|y_1 - z_1|, \ldots, |y_m - z_m|\}$ for m-dimensional vectors y, z, f, etc.

It is well known that when f is continuous on $D_f = [x_o, a] \otimes B_1 \otimes B_2 \ldots \otimes B_m$ with $a > x_o$ and y_{jo} in the interior of $B_j \in \mathcal{J}$ $(j = 1, \ldots, m)$, and when f satisfies a Lipschitz condition on D_f

$$(6.4) \qquad |f(x, y_1) - f(x, y_2)| \le K_f |y_1 - y_2|,$$

for some non-negative real number K_f, then there exists exactly one solution to (6.1) and (6.2) in $[x_o, x^*]$ for x^* such that, for all $(x,y) \in D_f$, we have

$$y_{jo} + (x^* - x_o)\, f_j(x,y) \in B_j \qquad (j = 1, 2, \dots, m) .$$

We will suppose throughout this section that $F_1, \dots, F_m$ are interval valued functions on the regular domain $D_F = \mathcal{I}_{[x_o,a]} \otimes \mathcal{I}_{B_1} \otimes \dots \otimes \mathcal{I}_{B_m}$, satisfying the following conditions for $j = 1, 2, \dots, m$:

1) F_j is continuous, and F_j restricted to $D_f = [x_o, a] \otimes B_1, \dots, \otimes B_m$ is a real valued function f_j, i.e., $F_j(x, y_1, \dots, y_m) = f_j(x, y_1, \dots, y_m)$ for $(x, y_1, \dots, y_m) \in D_f$;

2) F_j is inclusion monotonic, i.e., $X' \subset X$, $Y_1' \subset Y_1, \dots, Y_m' \subset Y_m$ implies

$$F_j(X', Y_1', \dots, Y_m') \subset F_j(X, Y_1, \dots, Y_m) ;$$

3) There is a real number K_F such that $w(F_j(X, Y_1, \dots, Y_m)) \leq K_F \max\{w(x), w(Y_1), \dots, w(Y_m)\}$.

Notice that in case F is a rational interval function on D_F with real restriction f on D_f then the conditions 1), 2), 3) are satisfied by F and also by $\overline{f}$, the united extension of f.

The conditions 1), 2), 3) above imply (6.4). To see this, let $Y_1 = [y_{11}, y_{12}], \ldots, Y_m = [y_{m1}, y_{m2}]$ or in abbreviated form $Y = [y_1, y_2]$. Assume $w(F_j(X, Y)) \leq K_F \max \{w(X), w(Y)\}$. Then $w(F_j(x, Y)) \leq K_F\, w(Y)$ for real $x \in [x_o, a]$. Since f_j is, by definition, the real restriction of F_j, we have $f_j(x,y) \in F_j(x, Y)$ whenever $y \in Y$. Therefore $f_j(x, y_1) - f_j(x, y_2) \in F_j(x, Y) - F_j(x, Y)$. Now

$$[a,b] - [a,b] = [-1, 1]\, w([a,b]) ,$$

so

$$|f_j(x, y_1) - f_j(x, y_2)| \leq w(F_j(x, Y)) \leq K_F |y_2 - y_1|$$

and therefore

$$|f(x, y_1) - f(x, y_2)| = \max_j |f_j(x, y_1) - f_j(x, y_2)| \leq K_F |y_2 - y_1| .$$

We notice incidentally that K_F serves as a Lipschitz constant for f. We conclude that conditions 1), 2), 3) guarantee the existence and uniqueness in $[x_o, x^*]$ of a solution to (6.1), (6.2) when f_j is the real restriction of F_j.

If $y_{jo} \in Y_{jo}$ for Y_{jo} properly contained in B_j then the equation

$$Y_{jo} + (X_j^* - x_o)\, F_j(D_f) = B_j$$

has a solution X_j^* with $w(X_j^*) > 0$ for each $j = 1, 2, \ldots, m$. By $F_j(D_f)$ we mean, of course, $F_j(D_f) = F_j([x_o, a], B_1, \ldots, B_m)$. Define

$$X^* = [x_o, a] \cap X_1^* \cap \cdots \cap X_m^* .$$

In this way we can compute an interval, namely X^*, in which existence and uniqueness of a solution y to (6.3), (6.2), is guaranteed.

The First Order Method.

Let n be a positive integer and define $X_i^{(n)}$, $y_{ji}^{(n)}$, $b_{ji}^{(n)}$ by $y_{jo}^{(n)} = y_{jo}$ and, for $i = 1, 2, \ldots, n$,

$$(6.5) \quad \begin{cases} X_i^{(n)} = \left[x_{i-1}^{(n)}, x_i^{(n)} \right] = x_o + [i-1, i] \dfrac{w(X^*)}{n} \\ b_{ji}^{(n)} = y_{j,i-1}^{(n)} + [0,1] \dfrac{w(X^*)}{n} F_j(D_f) \\ y_{ji}^{(n)} = y_{j,i-1}^{(n)} + \dfrac{w(X^*)}{n} F_j\left(X_i^{(n)}, b_{1i}^{(n)}, \ldots, b_{mi}^{(n)}\right) \end{cases}$$

In vector notation, dropping the superscript (n), writing h for $\frac{w(X^*)}{n}$, and writing $S = [0,1]h\, F(D_f)$, we simplify the writing of (6.5) to

$$X_i = x_o + [i-1, i]\, h ,$$

$$b_i = y_{i-1} + S ,$$

$$y_i = y_{i-1} + h\, F(X_i, b_i) ,$$

so that for $i = 1, 2, \ldots, n$, we have

$$y_i = y_{i-1} + h\, F(X_i, y_{i-1} + S) .$$

This recursion formula expresses, in its simplest form, our <u>first order</u> interval method for ordinary differential equations.

The solution y to (6.3), (6.2) clearly satisfies $y(x) \in y(x_{i-1}) + S$ for $x \in X_i$ (that is, for each $j = 1, 2, \ldots, m$, $y_j(x) \in y_j(x_{i-1}) + S_j$, etc.), and if $y(x_{i-1}) \in y_{i-1}$, then

$$y(x) = y(x_{i-1}) + \int_{x_{i-1}}^{x} f(x', y(x'))dx \qquad (x \in X_i)$$

so $y(x) \in y_{i-1} + (x - x_{i-1})\, F(X_i, y_{i-1} + S)$ whenever $x \in X_i$. Furthermore, writing $w(y_i) = \max w(y_{ji})$, etc., we find that

$$w(y_i) \leq w(y_{i-1}) + h\, K_F \max \{h, w(y_{i-1}) + c\, h\}$$

where $c = w([0,1]\, F(D_f))$; therefore

(6.6) $$w(y_i) \leq (\max (c, 1)) \left(e^{K_F(x_i - x_o)} - 1 \right) h .$$

Replacing the superscript, (n), we define for $n = 1, 2, \ldots$ the functions $y^{(n)}$ for all $x \in X^*$, noticing that $X^* = \bigcup_{i=1}^{n} X_i^{(n)}$, and difining $y^{(n)}(x)$ for $x \in X_i^{(n)}$,

(6.7) $$y^{(n)}(x) = y_{i-1}^{(n)} + \left(x - x_{i-1}^{(n)}\right) F\left(X_i, y_{i-1}^{(n)} + \left(X_i^{(n)} - x_{i-1}^{(n)}\right) F(D_f)\right).$$

The functions $y^{(n)}(x)$ are well defined since at $x_i^{(n)}$, the common end point of $X_i^{(n)}$ and $X_{i+1}^{(n)}$, we have

$$y_{i-1}^{(n)} + (x_i^{(n)} - x_{i-1}^{(n)})\ F\left(X_i,\ y_{i-1}^{(n)} + (X_i^{(n)} - x_{i-1}^{(n)})\ F(D_f)\right) \ = \ y_i^{(n)} .$$

In fact, the functions defined by (6.7) are obviously continuous interval valued functions and are piecewise linear in x; that is, for $0 \leq t \leq 1$ we can write

$$y^{(n)}((1 - t)x_{i-1} + t\ x_i) \ = \ (1 - t)y_{i-1}^{(n)} + t\ y_i^{(n)} .$$

(However, since $y_{i-1}^{(n)}$ is an interval we do <u>not</u> have $(1 - t)y_{i-1}^{(n)} = y_{i-1}^{(n)} - t\ y_{i-1}^{(n)}$, for $t > 0$.)

We have shown that the interval valued functions $y_j^{(n)}(x)$ defined by (6.5) and (6.7) contain the corresponding components of the solution to (6.1), (6.2); that is, for $n = 1, 2, \ldots$, and for $j = 1, 2, \ldots, m$ we have, recalling (6.6),

<u>Theorem</u> 6.8:

$$y_j(x) \ \in \ y_j^{(n)}(x) \qquad \text{for} \qquad x \in X^* ,$$

and the sequence of interval vector valued functions $y^{(1)}(x)$, $y^{(2)}(x)$, $y^{(3)}(x)$, ... converges uniformly to $y(x)$ for $x \in X^*$. Furthermore, there is a real number K such that for $x \in X^*$

$$\max_j w(y_j^{(n)}(x)) \ \leq \ \frac{K}{n} , \qquad j = 1, 2, \ldots, m . \tag{6.8}$$

The following example will serve to illustrate both the geometric and the computational significance of the above result.

Consider the equation

$$\frac{dy}{dx} = y^2 ,$$

and the initial condition

$$y(0) = 1 .$$

The rational interval function $G : \mathcal{I} \to \mathcal{I}$ defined by $G(Y) = Y^2$ has real restriction $G([y,y]) = f(y) = y^2$. In order to use the same notation as developed for the general case, we define $F : \mathcal{I}_{[0,a]} \otimes \mathcal{I}_B \to \mathcal{I}$ by $F(X,Y) = G(Y)$ so that $D_F = \mathcal{I}_{[0,a]} \otimes \mathcal{I}_B$ and F restricted to $[0,a] \otimes B$ is f and $D_f = [0,a] \otimes B$. We assume B is an interval of positive width containing the initial value $y(0) = 1$ in its interior, e.g., $1 \in [0,2]^o$, and that $a > 0$. The function F clearly satisfies conditions 1), 2), and 3).

Now $F(D_f) = F([0,a], B) = B^2$. Call X_1^* the solution of (compare (2.9))

$$1 + (X_1^*)B^2 = B .$$

Since $y(0) \in B$, we will always have $0 \in X_1^*$. Set $X^* = [0,a] \cap X_1^*$. Then we can be sure of the existence and uniqueness of a solution for $x \in X^*$. Figure 8 illustrates the geometric significance of the process for determining X^*.

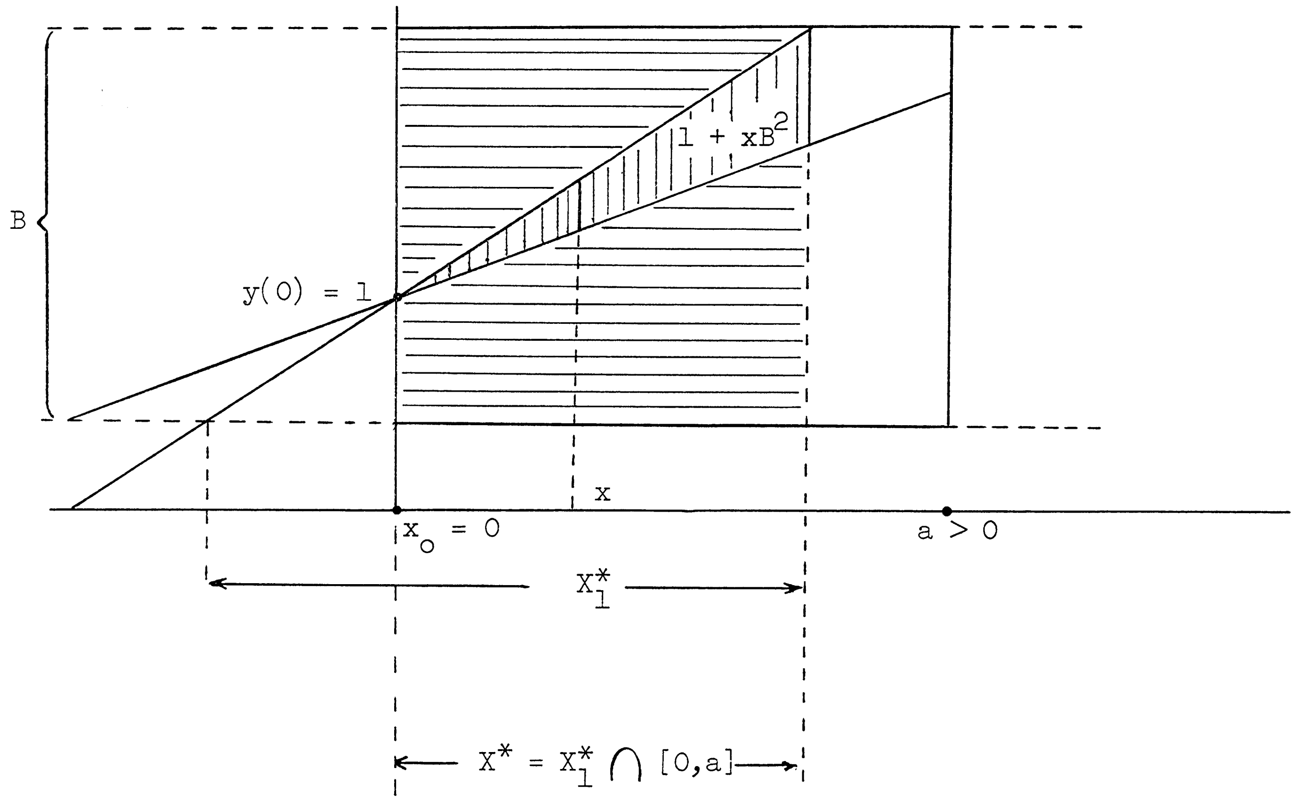

Figure 8

The shaded rectangle is $X^* \otimes B$. Let $B = [1/3, 2]$ and $a = 1$, for example; then the lines $y = 1 + (1/9)x$ and $y = 1 + 4x$ bound a wedge in $X^* \otimes B$ containing the solution $y(x)$ of $y' = y^2$, $y(0) = 1$; that is, for $x \in X^* = [0, 1/4]$ we have $y(x) \in 1 + xB^2 = 1 + x[1/9, 4]$, since in this case we have

$$1 + X_1^* [1/9, 4] = [1/3, 2] ,$$

$$X_1^* [1/9, 4] = [-2/3, 1] ,$$

or if $X^* = [c,d]$, then since $0 \in X^*$, we have $c \leq 0 \leq d$ and

$$[c,d][1/9,\ 4] = [4c,\ 4d]\ ,$$

so we find $4c = -2/3$, $4d = 1$ or $X_1^* = [-1/6,\ 1/4]$ and

$$X^* = [0,a] \cap X_1^* = [0,1] \cap [-1/6,\ 1/4] = [0,\ 1/4]\ .$$

Next, we determine for this example the functions $y^{(n)}$, defined by (6.5) and (6.7). We find that $w(X^*) = 1/4$ so $h = \frac{1}{4n}$ and, using $B = [1/3,\ 2]$, $a = 1$, we determine that

$$F(D_f) = B^2 = [1/9,\ 4]\ ,$$

$$S = [0,1]h\,B^2 = [0,1]\,\frac{1}{n}$$

$$X_i^{(n)} = [i-1,\ i]\,\frac{1}{4n} = \left[x_{i-1}^{(n)},\ x_i^{(n)}\right]\ ,$$

$$b_i^{(n)} = y_{i-1}^{(n)} + [0,1]\,\frac{1}{n}\ ,$$

$$y_i^{(n)} = y_{i-1}^{(n)} + \frac{1}{4n}\,b_i^2\ ,$$

and therefore

$$(6.9) \qquad y_i^{(n)} = y_{i-1}^{(n)} + \frac{1}{4n}\left(y_{i-1}^{(n)} + [0,1]\,\frac{1}{n}\right)^2\ ,$$

for $i = 1, 2, \ldots, n$, with $y_0^{(n)} = 1$.

The intervals $y_i^{(n)}$, $i = 1, 2, \ldots, n$, can be computed using (6.9) to obtain the functions $y^{(n)}(x)$. According to (6.7), we have

$$y^{(n)}(x) = y_{i-1}^{(n)} + \left(x - \frac{(i-1)}{4n}\right)\left(y_{i-1}^{(n)} + [0,1]\,\frac{1}{n}\right)^2 ,$$

(6.10)

$$\text{for } x \in [i-1, i]\,\frac{1}{4n} .$$

Evaluating (6.6), we find that

$$c = w([0,1]\,F(D_f)) = w([0,1]\,[1/9, 4])$$

$$= w([0,4]) = 4 ,$$

and if $Y = [y_1, y_2] \subset B = [1/3, 2]$, then

$$w(F(X, Y)) = w(Y^2) = w([y_1, y_2]^2)$$

$$= w([y_1^2, y_2^2]) = (y_2 - y_1)(y_2 + y_1)$$

$$= (y_1 + y_2)\,w(Y) ,$$

and we can take $K_F = 4$ in order to obtain $w(F(X,Y)) \leq K_F\,w(Y)$ for $(X,Y) \in D_F = \mathcal{I}_{[0,1]} \otimes \mathcal{I}_{[1/3,2]}$. Making these substitutions in (6.6), we obtain

(6.11) $$w(y_i^{(n)}) \leq \frac{1}{n}\,e^{\frac{i}{n}} , \qquad (i = 1, 2, \ldots, n) .$$

Figure 9 illustrates the construction of the functions $y^{(n)}(x)$ geometrically. The small rectangles are the b_i for $i = 1, 2, \ldots, n$, while the small dotted triangles are translates of $y_o + (x - x_o)\,F(D_f)$

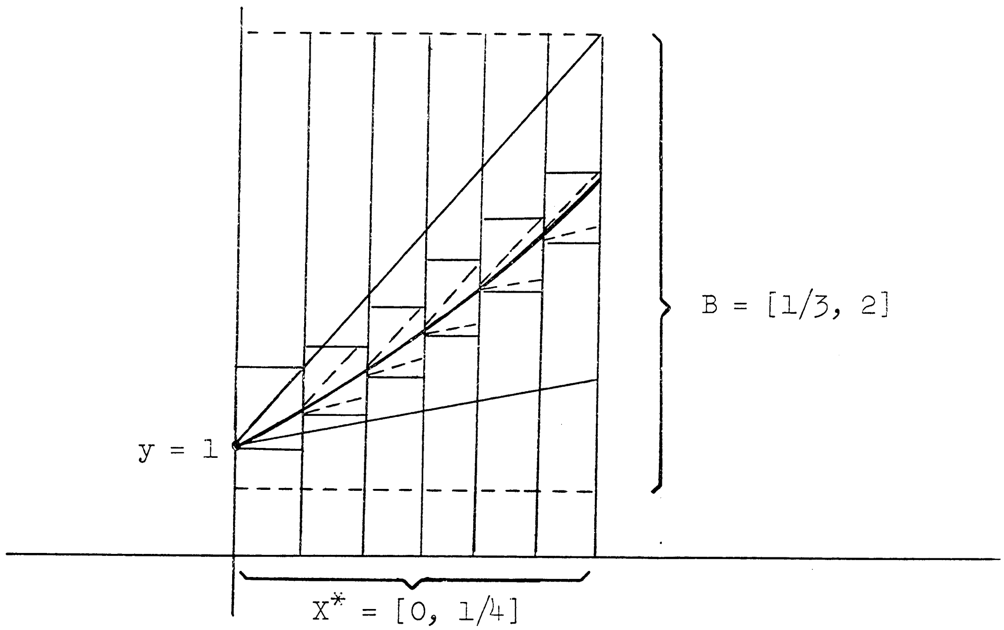

Figure 9

to the intervals y_{i-1}; that is, they represent the interval valued functions of $x \in X_i$ bounding the solutions y to $\frac{dy}{dx} = f(x,y)$ which pass through x_{i-1}, y_{i-1}. If $y(x_{i-1}) \in y_{i-1}$, then $y(x) \in y_{i-1} + (x - x_{i-1})\, F(D_f)$ for $x \in X_i$.

The choice of the interval B is seen to affect the width of X^* and the numbers c and K_F in (6.6). The bound expressed by (6.6) on the size of $w(y_i^{(n)})$ was derived in order to prove the convergence of the functions given by (6.7) to the solution of the differential equation. On the other hand, we only need the function F in order to compute the $y_{i-1}^{(n)}$ which determine (6.7) and we will automatically have $y(x) \in y^{(n)}(x)$; so that the interval valued function $y^{(n)}(x)$ gives upper and lower bounds to the solution $y(x)$ at each $x \in X^*$. For example, setting $n = 10$

in (6.9), we find by interval arithmetic computation that $y_{10}^{(10)} =$ $[1.321\ldots, 1.399\ldots]$ so $w(y_{10}^{(10)}) = .078\ldots$ which is about $1/4$ as big as $(1/10)e$ (compare (6.11)). The exact solution to $\frac{dy}{dx} = y^2$ with $y(0) = 1$ is, of course, given by $y(x) = \frac{1}{1 - x}$ and from (6.10), setting $i = 11$, we find that $y^{(10)}(1/4) = y_{10}^{(10)} = [1.321\ldots, 1.399\ldots]$. Thus, $y(1/4) = 4/3 = 1.33\ldots \in y^{(10)}(1/4)$ as promised.

Now having computed $y_i^{(n)}$ for $i = 1, 2, \ldots, n$ we can choose $y_n^{(n)}$ as a new initial condition at $x = x_n$, the right hand end point of the interval X^*. Select a new interval B containing $y_n^{(n)}$ in its interior and a new real number a, or perhaps use the same $a - x_o$ as before and set $B_{new} = y_n^{(n)} - y_o + B_{old}$. If the rectangle so determined still lies in the domain of F, we can proceed as before. In this way we can construct continuations of the functions $y^{(n)}(x)$. To illustrate, we will extend the $y^{(n)}(x)$ obtained for the example we have just treated above. See Figure 10.

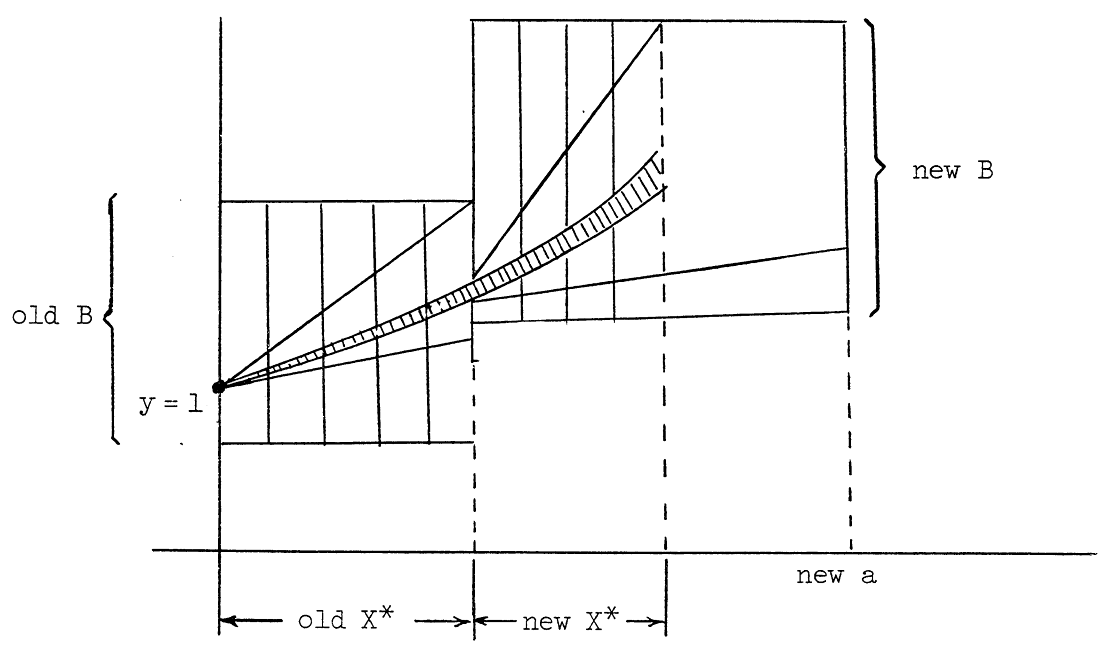

Figure 10

If we choose the new B so that $B = [1,d]$, with $d > y^{(n)}(1/4)$, then $F([1/4, a], B) = B^2 = [1, d^2]$ and the new X^* is determined by

$$y^{(n)}(\tfrac{1}{4}) + \left(X_1^* - \tfrac{1}{4}\right)[1, d^2] = [1,d] ,$$

and therefore

$$X^* = X_1^* \bigcap \left[\tfrac{1}{4} , a\right] = \left[\tfrac{1}{4} , a^*\right] .$$

If $y^{(n)}(\frac{1}{4}) = [y_1, y_2]$, then a^* is found from

$$y_2 + (a^* - \tfrac{1}{4})\, d^2 = d , \qquad\qquad d > y_2 ,$$

or

$$a^* = \frac{d - y_2}{d^2} + \frac{1}{4} .$$

Clearly, $d = 2y_2$ maximizes a^*; in fact,

$$a^* \leq \frac{1}{4} + \frac{1}{4y_2} .$$

Since $y_2 > \frac{4}{3}$, then $a^* < \frac{7}{16}$. We can again choose a positive integer n and compute the intervals $y_i^{(n)}$ bounding the solution over the new X^*. In fact, for any $0 < \delta < 1$ we can find a finite monotonic sequence of consecutive a^*'s such that the last one is in the interval $[1 - \delta, 1]$ by a finite repetition of the extension procedure. Then the constructed bounding functions $y^{(n)}(x)$ will converge uniformly on $[0, 1 - \delta]$ with increasing n to the solution $y(x) = \frac{1}{1 - x}$ and for each n and each $x \in [0, 1 - \delta]$ we will have $y(x) \in y^{(n)}(x)$.

The method we have been discussing is a first order method in the sense of (6.8); that is, the widths of the intervals $y^{(n)}(x)$ for fixed x are $O(n^{-1})$. It should be clear that in the example above, applying the method to the equation $y' = y^2$, the intervals $y^{(n)}(x)$ do <u>not</u> satisfy $w(y^{(n)}(x)) = o(n^{-1})$ for fixed $x > 0$. We will now turn to the investigation of a class of methods such that for each positive integer k, there is a <u>k^{th} order method</u> for constructing interval functions $y^{(k,n)}(x)$ of the real variable x (for x in an interval X^*) which are related to a solution y of (6.1), (6.2), by

$$y(x) \in y^{(k,n)}(x) \qquad \text{for} \qquad x \in X^*$$

and such that $w(y^{(k,n)}(x)) = O(n^{-k})$; in fact, for each k there will be a positive real number K_k such that for all $x \in X^*$ and for all positive integers n,

$$w(y^{(k,n)}(x)) \leq K_k(\tfrac{1}{n})^k .$$

<u>The Methods of Order</u> $k > 1$.

We will derive a k^{th} order interval method for each integer $k > 1$ based on local expansion in a Taylor series in a fashion similar to the development of $I_{n,k}$ in Section 5.

To avoid notational complication we will assume k is a fixed integer, $k > 1$, in the rest of this section.

Let $F = (F_1, \ldots, F_m)$ satisfy conditions 1), 2), 3) stated in the first part of this section. Furthermore, let $F_j^{(\ell)}$, $j = 1, 2, \ldots, m$; $\ell = 0, 1, \ldots, k-1$, be interval valued functions also defined on D_F with real valued real restrictions $f_j^{(\ell)}$ such that $f_j^{(\ell)} = \frac{d}{dx} f_j^{(\ell-1)}$ on D_f; that is,

$$f_j^{(\ell)} = \frac{\partial f_j^{(\ell-1)}}{\partial x} + \sum_{r=1}^{m} \frac{\partial f_j^{(\ell-1)}}{\partial y_r} f_r \qquad (\ell = 1, 2, \ldots, k-1),$$

with $f_j^{(0)} = f_j$.

Assume $K_F^{(\ell)}$, $\ell = 0, 1, \ldots, k-1$ are positive real numbers such that $F_j^{(\ell)}$ satisfies conditions 2), 3), with

$$(6.12) \quad w(F_j^{(\ell)}(X, Y_1, \ldots, Y_m)) \leq K_F^{(\ell)} \max (w(X), w(Y_1), \ldots, w(Y_m)) .$$

For example, if $F_j^{(\ell)}$, $j = 1, 2, \ldots, m$; $\ell = 0, 1, \ldots, k-1$, are rational interval functions on D_F then all these conditions are satisfied if $F_j^{(\ell)}$ has real restriction $f_j^{(\ell)}$, a real rational function on D_f, satisfying $f_j^{(\ell)} = \frac{d}{dx} f_j^{(\ell-1)}$.

Using vector notation again we have $F^{(\ell)} = (F_1^{(\ell)}, \ldots, F_m^{(\ell)})$, $f^{(\ell)} = (f^{(\ell)}, \ldots, f_m^{(\ell)})$, etc.

We define the function $A(X, x, Y) = (A_1(X, x, Y), \ldots, A_m(X, x, Y))$ on $\mathcal{I}_{[x_o, a]} \otimes D_F$ with $Y = (Y_1, \ldots, Y_m)$ by

$$A(X, x, Y) = Y + \sum_{\ell=1}^{k-1} \frac{F^{(\ell-1)}(x, Y)}{\ell!} (X - x)^{\ell} + \frac{F^{(k-1)}(D_f)}{k!} (X - x)^k \quad . \tag{6.13}$$

The function A will play a role similar to that of $1 + xB^2$ in Figure 8 and the dotted triangles of Figure 9.

Let $B = (B_1, \ldots, B_m)$ and $C = (C_1, \ldots, C_m)$. Then by $C \subset B$ we mean $C_j \subset B_j$, $j = 1, \ldots, m$.

Recall that $D_F = \mathcal{J}_{[x,a]} \otimes \mathcal{J}_{B_1}, \ldots, \otimes \mathcal{J}_{B_m}$ with y_{jo} in the interior of B_j, $j = 1, \ldots, m$. We wish to determine an X^* such that $w(X^*) > 0$ and $X^* \subset [x_o, a]$ and $A(X^*, x_o, y_o) \subset B$. The widest X^* satisfying these conditions is determined by the equations

$$A_j(X^*, x_o, y_o) = B_j , \qquad j = 1, \ldots, m$$

(which are not linear in X^* for $k > 1$). Rather than assume a solution for X^* we proceed as follows in order to determine a suitable X^* by a finite sequence of evaluations of A. Choose a positive integer p. Compute $A^{(o)} = A([x_o, a], x_o, y_o)$. If $A^{(o)} \subset B$ then set $X^* = [x_o, a]$, otherwise find the smallest positive integer q, such that

$$A^{(q)} = A([x_o, x_o + 2^{-q}(a - x_o)], x_o, y_o) \subset B .$$

This is done by computing $A^{(1)}, A^{(2)}, \ldots$ successively until the first q for which $A^{(q)} \subset B$. Such a q exists by the assumption that y_{jo}

is in the interior of B_j, $j = 1, 2, \dots, m$. If $p = 1$, then set $X^* = [x_o, x_o + 2^{-q}(a - x_o)]$. To get a slightly wider X^* we can choose $p > 1$ and determine successively $p - 1$ binary digits, $b_2, \dots, b_p$ (i.e., each $b_s = 0$ or 1, $s = 2, \dots, p$) such that for $\theta = 1 \cdot 2^{-q} + b_2 \cdot 2^{-q-1} + \dots + b_p 2^{-q-p+1}$, we have

$$A([x_o, x_o + \theta(a - x_o)], x_o, y_o) \subset B .$$

For every positive integer n, define

$$h_n = \frac{w(X^*)}{n} ,$$

$$X_i^{(n)} = x_o + [i - 1, i]h_n = \left[x_{i-1}^{(n)}, x_i^{(n)}\right] , \qquad i = 1, 2, \dots, n$$

so that $x_o^{(n)} = x_o$ and $x_n^{(n)} = x_o + w(X^*)$ for all n. Define $y_o^{(n)} = y_o$ for all n and for $i = 1, 2, \dots, n$; define for $x \in X_i^{(n)}$,

$$(6.14)\begin{cases} y_i^{(n)}(x) = A(x, x_o, y_o) \\ \qquad \bigcap \left\{ y_{i-1}^{(n)} + \sum_{\ell=1}^{k-1} \frac{F^{(\ell-1)}(x_{i-1}^{(n)}, y_{i-1}^{(n)})}{\ell!} (x - x_{i-1}^{(n)})^\ell \right. \\ \qquad \left. + \frac{F^{(k-1)}\left(X_i^{(n)}, A(X_i^{(n)}, x_{i-1}^{(n)}, y_{i-1}^{(n)})\right)}{k!} (x - x_{i-1}^{(n)})^k \right\} \end{cases}$$

then

$$y_i^{(n)}(x_{i-1}^{(n)}) = y_{i-1}^{(n)} \bigcap A(x_{i-1}, x_o, y_o) ,$$

so call $y_i^{(n)}(x_i^{(n)}) = y_i^{(n)}$ for $i = 1, 2, \ldots, n$ and (6.14) defines, by finite induction on i, a function $y^{(n)}$ on $x \in X^*$ for each n, by $y^{(n)}(x) = y_i^{(n)}(x)$ for $x \in X_i^{(n)}$. The quantities $y_i^{(n)}$, $i = 1, \ldots, n$ are each determined by a finite number of evaluations of the $F^{(\ell)}$; substituting $x_i^{(n)}$ for x on the right hand side of (6.14), the left hand side becomes $y_i^{(n)}(x_i^{(n)}) = y_i^{(n)}$.

Recall that we are using vector notation, so that the quantity $y_i^{(n)}$, for example, is the m-tuple of intervals $(y_{1i}^{(n)}, \ldots, y_{mi}^{(n)})$. And $w(y_i^{(n)}) = \max_j w(y_{ji}^{(n)})$, etc.

It is not hard to show that there is a positive real number M, such that for all positive integers n and for all $x \in X^*$,

$$(6.15) \qquad w(y^{(n)}(x)) \leq M\left(\frac{w(X^*)}{n}\right)^k .$$

From (6.12), (6.13), (6.14) one derives an inequality of the form $w(y_i^{(n)}) \leq (1 + h_n K)\, w(y_{i-1}^{(n)}) + c\, h_n^{k+1}$. Then (6.15) follows easily. In the present situation $w(y_o^{(n)}) = 0$. More generally, (6.15) holds provided $w(y_o^{(n)}) \leq N(\frac{1}{n})^k$ for some $N > 0$. This fact will permit the continuation of the functions $y^{(n)}$ to a new X^* without losing the inequality (6.15) on the union of the new X^* and the old X^*.

Instead of assuming a solution y to (6.1), (6.2) we will give a proof of the following assertion:

<u>Theorem</u> 6.16:

The equation

$$y(x) = \bigcap_{n=1}^{\infty} y^{(n)}(x) , \qquad x \in X^* ,$$

defines a function $y \in C^k(X^*)$ satisfying (6.1), (6.2).

Proof:

First, we wish to show that for every $x \in X^*$ and for each pair of positive integers n_1, n_2, the interval valued functions $y^{(n_1)}$, $y^{(n_2)}$, defined above, have non-empty intersection at x; i.e., $y^{(n_1)}(x) \cap y^{(n_2)}(x)$ is non-empty.

In fact, we will use (6.15) to show that for some positive integer N_o and for every $x \in X^*$, $n > N_o$ implies

$$y^{(nn_1n_2)}(x) \subset y^{(n_1)}(x) \cap y^{(n_2)}(x) \;. \tag{6.16}$$

It is sufficient in order to demonstrate (6.16) to show that for any n_1 there is a large enough N_o such that $n' > N_o$ implies $y^{(n'n_1)}(x) \subset y^{(n_1)}(x)$ for all $x \in X^*$.

From the definition of $y^{(n)}(x)$ it is clear that $y^{(n)}(x) \subset A(x, x_o, y_o)$ for $x \in X^*$. Furthermore, $y^{(n)}(x)$ is non-empty in $X_1^{(n)}$. Since

$$A\left(x, x_{i-1}^{(n)}, y_{i-1}^{(n)}\right) = y_{i-1} + \sum_{\ell=1}^{k-1} \frac{F^{(\ell-1)}(x_{i-1}^{(n)}, y_{i-1}^{(n)})}{\ell!}\left(x - x_{i-1}^{(n)}\right)^k + \frac{F^{(k-1)}(D_f)}{k!}\left(x - x_{i-1}^{(n)}\right)^k$$

and

$$\left(X_1^{(n)},\ A(X_1^{(n)},\ x_o,\ y_o)\right) \subset D_f ,$$

it follows that the expression in brackets for $i = 2$ on the right hand side of (6.14),

$$y_1^{(n)} + \sum_{\ell=1}^{k-1} \frac{F^{(\ell-1)}\left(x_1^{(n)}, y_1^{(n)}\right)}{\ell!} \left(x - x_1^{(n)}\right)^{\ell}$$

$$+ \frac{F^{(k-1)}\left(X_2^{(n)}, A\left(X_2^{(n)}, x_1^{(n)}, y_1^{(n)}\right)\right)}{k!} \left(x - x_1^{(n)}\right)^{k}$$

is contained in $A(x, x_1^{(n)}, y_1^{(n)})$ for $x \in X_2^{(n)}$.

We claim that for n sufficiently large,

$$A\left(x, x_{i-1}^{(n)}, y_{i-1}^{(n)}\right) \subset A(x, x_o, y_o)$$

for $x \in X_i^{(n)}$, and therefore $y^{(n)}(x)$ is non- empty in $X_1^{(n)} \cup X_2^{(n)} \cup \dots \cup X_i^{(n)}$. Proceeding in this way for $i = 2, 3, \dots, n$, we finally have that $y^{(n)}(x)$ is non-empty in X^*.

Imitating the above argument with $y^{(n_1)}(x) \cap y^{(n'n_1)}(x)$ in place of $A(x, x_o, y_o) \cap \{ \cdots \}$ on the right hand side of (6.14), we have for sufficiently large N_o that $n' > N_o$ implies

$$y^{(n'n_1)}(x) \subset y^{(n_1)}(x), \quad \text{for all} \quad x \in X^* .$$

Thus, for each $x \in X^*$ and every finite collection of positive integers $n_1, n_2, \dots, n_p$, we have that

$$\bigcap_{q=1}^{p} y^{(n_q)}(x)$$

is non-empty and contained in B. By the finite intersection property of the compact set B this means we can define a function y on X^* by

$$y(x) = \bigcap_{n=1}^{\infty} y^{(n)}(x) .$$

For each $x \in X^*$, $y(x)$ is non-empty; in fact, it is clearly a real m-tuple. By (6.15),

$$\max_j w(y_j(x)) = w(y(x)) \leq M \frac{w(X^*)}{n}^k , \qquad j = 1 , 2 , \dots , m ,$$

for all n; hence, $w(y(x)) = 0$ for all $x \in X^*$, that is, $y_j(x) = [y_j(x), y_j(x)]$ or $y_j(x)$ is a real number. Notice that $y(x) \in y^{(n)}(x)$ for all n and all $x \in X^*$.

We claim that y satisfies (6.3) with $y(x_o) = y_o$, that is, $y(x) = (y_1(x), \dots , y_m(x))$ with $\{y_j\}$ satisfying (6.1) and (6.2).

For large enough n we have, for $i = 1, 2, \dots , n$,

$$(6.17) \qquad \begin{aligned} y_i^{(n)} &= y_{i-1}^{(n)} + \sum_{\ell=1}^{k-1} \frac{F^{(\ell-1)}\left(x_{i-1}^{(n)} , y_{i-1}^{(n)}\right)}{\ell!} h_n^{\ell} \\ &+ \frac{F^{(k-1)}\left(X_i^{(m)}, A(X_i^{(n)}, x_{i-1}^{(n)} , y_{i-1}^{(n)})\right)}{k!} h_n^k , \end{aligned}$$

where $h_n = \frac{w(X^*)}{n}$, as before.

Recall that $x_{i-1}^{(n)} = x_o + (i - 1)h_n$. Now fix $x_{i-1}^{(n)}$, then $x_i^{(n)} = x_{i-1}^{(n)} + h_n$ and let n be large enough so that (6.17) holds, then

$$\left| \frac{y\left(x_{i-1}^{(n)} + h_n\right) - y\left(x_{i-1}^{(n)}\right)}{h_n} - f\left(x_{i-1}^{(n)}, y\left(x_{i-1}^{(n)}\right)\right) \right|$$

$$(6.18) \qquad \le \frac{1}{h_n} \left\{ w\left(y_{i-1}^{(n)} + h_n F^{(o)}\left(x_{i-1}^{(n)}, y_{i-1}^{(n)}\right)\right) + w(y_i^n) + \left| h_n^2 \frac{F^{(1)}}{2!} + \cdots + h_n^k \frac{F^{(k-1)}}{k!} \right| \right\} .$$

Since $k > 1$, then by (6.15) and (6.18) we have

$$\frac{dy}{dx}\left(x_{i-1}^{(n)}\right) = f\left(x_{i-1}^{(n)}, y\left(x_{i-1}^{(n)}\right)\right) .$$

In fact, with a little refinement of the above argument it follows that for $x \in X^*$, $\frac{dy}{dx}(x) = f(x, y(x))$ and $y \in C^k(X^*)$. The differentiability of y follows from that of f.

We will now illustrate the computational aspects of the k^{th} order procedure for $k > 1$ with an example. We consider the same example used to illustrate the first order method above, namely

$$\frac{dy}{dx} = y^2 = f(x,y) , \quad y(o) = 1 .$$

Thus, let

$$F^{(0)}(X, Y) = Y^2 ,$$

$$F^{(1)}(X, Y) = 2Y^3 ,$$

$$\cdots$$

$$F^{(k-1)}(X, Y) = k!\, Y^{k+1} ,$$

and let $D_f = [0,1] \otimes [1/3, 2]$, as before. From (6.13) we find that

$$A(X, x, Y) = Y + \sum_{p=1}^{k-1} Y^{\ell+1}(X - x)^{\ell} + [1/3, 2]^{k+1} (X - x)^k .$$

First we need to determine an X^* such that

$$w(X^*) > 0 ,$$

$$X^* \subset [0,1] ,$$

$$A(X^*, 0, 1) \subset [1/3, 2] .$$

Using the procedure described following (6.14) we find that

$$A([0,1], 0, 1) = 1 + \sum_{\ell=1}^{k-1} [0,1]^{\ell} + [1/3, 2]^{k+1} [0,1]^k \not\subset [1/3, 2] .$$

So we compute

$$A([0,1/2], 0, 1) = 1 + \sum_{\ell=1}^{k-1} [0,1/2]^{\ell} + [1/3,2]^{k+1} [0,1/2]^k \not\subset [1/3,2]$$

until the first q, for which

$$A([0, 2^{-q}], 0, 1) \subset [1/3, 2] .$$

It turns out in this example that

$$A([0,1/4], 0, 1) = \left[1, 1 + (\tfrac{1}{4}) + \cdots + (\tfrac{1}{4})^{k-1} + 2^{k+1} \cdot (\tfrac{1}{4})^{k}\right] \subset [1/3,2] ,$$

for all $k > 1$. So we can use $X^* = [0, 1/4]$. If $k = 4$, for example, we could also use $X^* = [0, 7/16]$. Setting $p = 3$, we find that for $k = 4$,

$$A\left(\left[0, 2^{-2} + b_2 \cdot 2^{-3} + b_3 \cdot 2^{-4}\right], 0, 1\right) \subset [1/3, 2] ,$$

with $b_2 = b_3 = 1$.

We will use $X^* = [0, 1/4]$ with all $k > 1$. Thus,

$$h_n = \frac{w(X^*)}{n} = \frac{1}{4n} ,$$

$$X_i^{(n)} = [i - 1, i] \frac{1}{4n} = \left[x_{i-1}^{(n)} , x_i^{(n)}\right] ,$$

so that

$$x_i^{(n)} = \frac{i}{4n} , \qquad i = 1 , 2 , \ldots , n .$$

Now

$$A\left(X_i^{(n)}, x_{i-1}^{(n)} , y_{i-1}^{(n)}\right) = y_{i-1}^{(n)} + \sum_{\ell=1}^{k-1} \left(y_{i-1}^{(n)}\right)^{\ell+1} \left[0, \frac{1}{4n}\right]^{\ell} + \left[0, \frac{2}{(2n)^k}\right]$$

and

$$A(x, 0, 1) = 1 + x + \cdots + x^{k-1} + [1/3, 2]^{k+1} x^k .$$

So from (6.14) for $x \in X_i^{(n)}$, $i = 1, 2, \ldots, n$ we have

$$\begin{aligned} y_i^{(n)}(x) = & \left\{1 + x + \cdots + x^{k-1} + \left[\frac{1}{3}, 2^{k+1}\right] x^k\right\} \\ & \cap \left\{ y_{i-1}^{(n)} + \sum_{\ell=1}^{k-1} \left(y_{i-1}^{(n)}\right)^{\ell+1} \left(x - x_{i-1}^{(n)}\right)^{\ell} \right. \\ & + \left(y_{i-1}^{(n)} + \sum_{\ell=1}^{k-1} \left(y_{i-1}^{(n)}\right)^{\ell+1} \left[0, \frac{1}{4n}\right]^{\ell} + \left[0, \frac{2}{(2n)^k}\right]\right)^{k+1} \\ & \left. \cdot \left(x - x_{i-1}^{(n)}\right)^k \right\} , \end{aligned} \tag{6.19}$$

with $x_o^{(n)} = 0$, $y_o^{(n)} = 1$, and $y_i^{(n)} = y_i^{(n)}(x_i^{(n)})$.

The results obtained assert that for each $k > 1$ there is an $M_k > 0$ such that for all $x \in [0, 1/4]$ and all n

$$y(x) = \bigcap_{n=1}^{\infty} y^{(n)}(x) = \frac{1}{1 - x} \in y^n(x)$$

and

$$w(y^{(n)}(x)) \leq M_k\left(\frac{1}{4n}\right)^k$$

with $y^{(n)}(x) = y_i^{(n)}(x)$ for $x \in X_i^{(n)} = [i - 1, i] \frac{1}{4n}$ where $y_i^{(n)}(x)$ is given by (6.19).

We will verify these assertions. Call $h = \frac{1}{4n}$. From (6.19) with $i = 1$, we have for $x \in X_1^{(n)}$,

$$(6.20) \quad y_i^{(n)}(x) = 1 + x + \cdots + x^{k-1} + [1, 1+h + \cdots + h^{k-1} + 2^{k+1} h^k]x^k .$$

We prove by induction on i that

$$\frac{1}{1-x} \in y_i^{(n)}(x) \quad \text{for} \quad x \in X_i^{(n)} .$$

For any $k > 1$, we have for $x \in X^* = [0, 1/4]$,

$$\frac{1}{1-x} = 1 + x + \cdots + x^{k-1} + \frac{x^k}{1-x} \in 1 + x + \cdots + x^{k-1} + [1, 4/3]\, x^k ,$$

therefore

$$\frac{1}{1-x} \in A(x, 0, 1) \quad \text{for} \quad x \in X^* .$$

Now for $x \in X_1^{(n)} = [0,h]$, we have

$$\frac{1}{1-x} = 1 + x + \cdots + x^{k-1} + \left(1 + x + \cdots + x^{k-1} + \frac{x^k}{1-x}\right) x^k ,$$

therefore

$$\frac{1}{1-x} \in 1 + x + \cdots + x^{k-1} + \left(1 + [0,h] + \cdots + [0,h]^{k-1} + \frac{[0,h]^k}{1-[0,h]}\right) x^k$$

since $n \geq 1$, we have

$$\frac{1}{1 - [0,h]} = \left[1, \frac{1}{1 - \frac{1}{4n}}\right] \subset [1, 4/3] .$$

By (6.20), we therefore have

$$\frac{1}{1 - x} \in y_1^{(n)}(x) \quad \text{for} \quad x \in X_1^{(n)} .$$

Assume for some $i \geq 2$, that

$$\frac{1}{1 - x} \in y_{i-1}^{(n)}(x) \quad \text{for} \quad x \in X_{i-1}^{(n)} .$$

We want to show that it follows that

$$\frac{1}{1 - x} \in y_i^{(n)}(x) \quad \text{for} \quad x \in X_i^{(n)} .$$

By the inductive hypothesis, we have

$$\frac{1}{1 - x_{i-1}^{(n)}} \in y_{i-1}^{(n)} = y_{i-1}^{(n)}\left(x_{i-1}^{(n)}\right) .$$

And we have already shown that

$$\frac{1}{1 - x} \in A(x, 0, 1) \quad \text{for} \quad x \in X^* .$$

Call $t = x - x_{i-1}^{(n)}$ and $u = \frac{1}{1 - x_{i-1}^{(n)}}$, then for $x \in X_i^{(n)} = x_{i-1}^{(n)} + [0,h]$, we have

$$\frac{1}{1-x} = \frac{1}{1 - x_{i-1}^{(n)} - t} = u\left(1 + tu + \cdots + (tu)^{k-1} + (tu)^k \frac{1}{1-tu}\right)$$

$$(6.21) \qquad = u + u^2 t + \cdots + u^k t^{k-1}$$

$$+ (ut)^k \left(u + u^2 t + \cdots + u^k t^{k-1} + \frac{(ut)^k}{1-x}\right) .$$

Now call

$$Z = y_{i-1}^{(n)} + \sum_{\ell=1}^{k-1} \left(y_{i-1}^{(n)}\right)^{\ell+1} \left[0 , \frac{1}{4n}\right]^{\ell} + \left[0 , \frac{2}{(2n)^k}\right] ,$$

then

$$ut = \frac{x - x_{i-1}^{(n)}}{1 - x_{i-1}^{(n)}} \in \left(y_{i-1}^{(n)}\right)\left(x - x_{i-1}^{(n)}\right) \subset (Z)\left(x - x_{i-1}^{(n)}\right)$$

and $\frac{1}{1-x} \in [0,2]$, so

$$\frac{(ut)^k}{1-x} \in \left[0, 2^{k+1} h^k\right] .$$

Putting these relations together, it follows from (6.21) and (6.19) that

$$\frac{1}{1-x} \in y_i^{(n)}(x) \quad \text{for} \quad x \in X_i^{(n)} .$$

Therefore, by induction on i, we have for $x \in X^*$,

$$\frac{1}{1-x} \in y^{(n)}(x) .$$

Next, we wish to show for $x \in X^*$ that

$$w\left(y^{(n)}(x)\right) \leq M_k \left(\frac{1}{4n}\right)^k \quad \text{for some} \quad M_k > 0 .$$

From (6.19) and the fact that $y_i^{(n)}(x) \subset [0,2]$, it follows that

$$\text{(6.22)} \quad w(y_i^{(n)}) \leq w(y_{i-1}^{(n)}) \left\{ 1 + \frac{1}{2n}\left(\sum_{\ell=1}^{k-1} (\ell + 1)(\frac{1}{2n})^{\ell-1} + \frac{2(k+1)}{(2n)^{k-1}} \sum_{\ell=1}^{k} \ell (\frac{1}{2n})^{\ell-1} \right) \right\} + 2(\frac{1}{2n})^{k+1} (k+1) \left(\frac{1}{(2n)^{k-1}} + \sum_{\ell=0}^{k} (\frac{1}{2n})^{\ell} \right)$$

It follows from this inequality for $x \in X^*$, that there is a K_k such that

$$\text{(6.23)} \qquad w(y^{(n)}(x)) \leq K_k \left(\frac{1}{4n}\right)^k$$

and in particular, for $n \geq k+1$, we find that

$$w(y^{(n)}(x)) \leq (k + 1)\, 2^{k+3} \left(\frac{1}{4n}\right)^k ,$$

and for large enough n, we find that

$$w(y^{(n)}(x)) \leq (k + 1)\, 2^{k+1} \left(\frac{1}{4n}\right)^k .$$

Since we have already shown that for all positive integers n and for all $x \in X^*$,

$$\frac{1}{1-x} \in y^{(n)}(x)$$

and since we now have $w(y^{(n)}(x)) \to 0$ with $n \to \infty$, we can conclude that

$$y(x) = \bigcap_{n=1}^{\infty} y^{(n)}(x) = \frac{1}{1-x} .$$

This, together with the inequality (6.23), completes the verification of the assertions made above.

The number of arithmetic operations required to evaluate $y_i^{(n)}$ for $i = 1, 2, \ldots, n$, is roughly nk^2. If we wish to minimize nk^2 by a choice of n, k such that

$$w(y^{(n)}(x)) \leq 10^{-10}$$

say, then assuming $w(y^{(n)}(x))$ is proportional to $(k+1)(\frac{1}{2n})^k$ we are led to the choice $k = 12$, $n = 5$ for which the number of operations is about 700, which must be reasonably close to the best choice. By direct verification using (6.22), we find that for $k = 12$ and $n = 5$,

$$w(y^{(5)}(x)) \leq \frac{1}{4} \cdot 10^{-10} .$$

Comparing this with the first order method, $k = 1$, discussed earlier we see that using (6.11) we would have to take $n \geq 10^{10} \cdot e$ to guarantee $w(y^{(n)}(x)) \leq 10^{-10}$ with the first order method. The number of arithmetic operations then would be (compare (6.10)) roughly 10^{11} which is, to say the least, much larger than the 700 required for $k=12$, $n=5$.

The methods of this section apply to any differential system (6.1) with $k - 1$ times differentiable functions f_j on a rectangular region D_f in $m + 1$ dimensional Euclidean space, taking in particular,

$$F_j^{(\ell)} = \overline{f}_j^{(\ell)} , \qquad \ell = 0 , 1 , \ldots , k - 1 ,$$

with $\overline{f}_j^{(\ell)}$ the united extension of $f_j^{(\ell)}$. This fact can be exploited computationally.

For example, in Section 5 we found interval valued functions I_n, such that $\{\ell n\, y \mid y \in Y'\} \subset I_n(Y)$. Recall that, for $Y > 1$,

$$I_N(Y) = \sum_{i=1}^{N} \frac{w\left([i - 1, i] \frac{Y - 1}{N}\right)}{1 + [i - 1, i] \frac{Y - 1}{N}}$$

and for $Y = [y_1, y_2]$, with $Y_j^{(M)} = y_1 + [j - 1, j] \frac{w(Y)}{M}$, we have

$$I_{N,M}(Y) = \bigcup_{j=1}^{M} I_N\left(Y_j^{(M)}\right) = \{\ell n\, y \mid y \in Y\} + E_{N,M} ,$$

with

$$0 \in E_{N,M} , \quad w(E_{N,M}) \leq \frac{K\, w(Y)}{M} + \frac{2(|Y| - 1)}{N} .$$

Let $\overline{\ell n}$ be the united extension of ℓn so that

$$\overline{\ell n}\, Y = \{\ell n\, y \mid y \in Y\} , \qquad Y > 1 .$$

Then $\overline{\ell n}\, Y \subset I_n(Y)$ and if the function ℓn appears in a system (6.1), for example,

$$(6.24) \qquad \begin{array}{ccc} y_1' & = & \ell n\, y_1 + g(x,y) \\ \vdots & \vdots & \vdots \end{array}$$

then computations paralleling the $\underline{k^{th}}$ order method with a particular choice of n for $F^{(0)}(X, Y) = \overline{\ell n}\, Y + G^{(0)}(X, Y)$, $F^{(1)}(X, Y) = \frac{1}{Y} + G^{(1)}(X, Y)$, etc., can be carried out (assuming D_f is such that $Y > 1$) instead with $F^{(0)}(X, Y) = I_{N,M}(Y) + G^{(0)}(X, Y)$, $F^{(1)}(X, Y) = \frac{1}{Y} + G^{(1)}(X,Y)$, etc., yielding functions $y^{(n)}$ bounding the solution to (6.24) for a particular initial value problem. With N, M large enough the widths of the bounding intervals $y_i^{(n)}$ to the solution can be made arbitrarily close to those which result from using $\overline{\ell n}\, Y$.

Alternatively, since the function ℓn itself satisfies a rational differential equation,

$$(\ell n\, y)' = \frac{1}{y} .$$

We can add a new equation to the system (6.24) of the form

$$\frac{d}{dx} y_{m+1} = \frac{1}{y_{m+1}}$$

and replace $\ell n\, y_1$ wherever it appears by y_{m+1}, using the initial condition

$$y_{m+1}(x_o) = \ell n\, x_o .$$

Similarly, other non-rational functions such as exp, $\sqrt{\ }$, sin, etc., can be handled in differential systems using the methods of this section either by replacing them directly by finitely computable interval functions, or by adding rational differential equations to the given system so that the augmented system is rational.

If a differential system (6.1) is given in which certain parameters $\lambda_1 , \dots , \lambda_p$ occur rationally we can find at once bounding intervals to the components of families of solutions corresponding to invervals $L_1 , \dots , L_p$ of values $\lambda_1 , \dots , \lambda_p$, simply by using interval coefficients $L_1, \dots , L_p$ in place of $\lambda_1, \dots , \lambda_p$ during the computations of the first or $\underline{k^{th}}$ order methods, $k > 1$, described above.

Also, bounds on the components of a family of solutions to a differential system corresponding to intervals of initial values $y_{jo} \in Y_{jo}$ can be found in a similar way.

The following remarks, however, indicate a computational limitation of such a scheme.

Consider the differential system

$$\frac{dy_1}{dx} = y_2 , \quad \frac{dy_2}{dx} = -y_1 ,$$

and suppose we wish to compute by a finite number of interval operations bounds on the set of values at each x in some interval of the family of solutions passing through

$$x_o = 0 , \quad y_{10} = 0 , \quad y_{20} \in Y_{20} = [1,2] .$$

Since the solution $y = (y_1, y_2)$, corresponding to a particular $y_{20} \in [1,2]$, is obviously given by $y(x) = (y_{20} \sin x, y_{20} \cos x)$, then the set of values of y at $x = \frac{\pi}{4}$ is the segment

$$T = \left\{ \left(\frac{t}{\sqrt{2}}, \frac{t}{\sqrt{2}} \right) \mid t \in [1,2] \right\}.$$

(See Figure 11.)

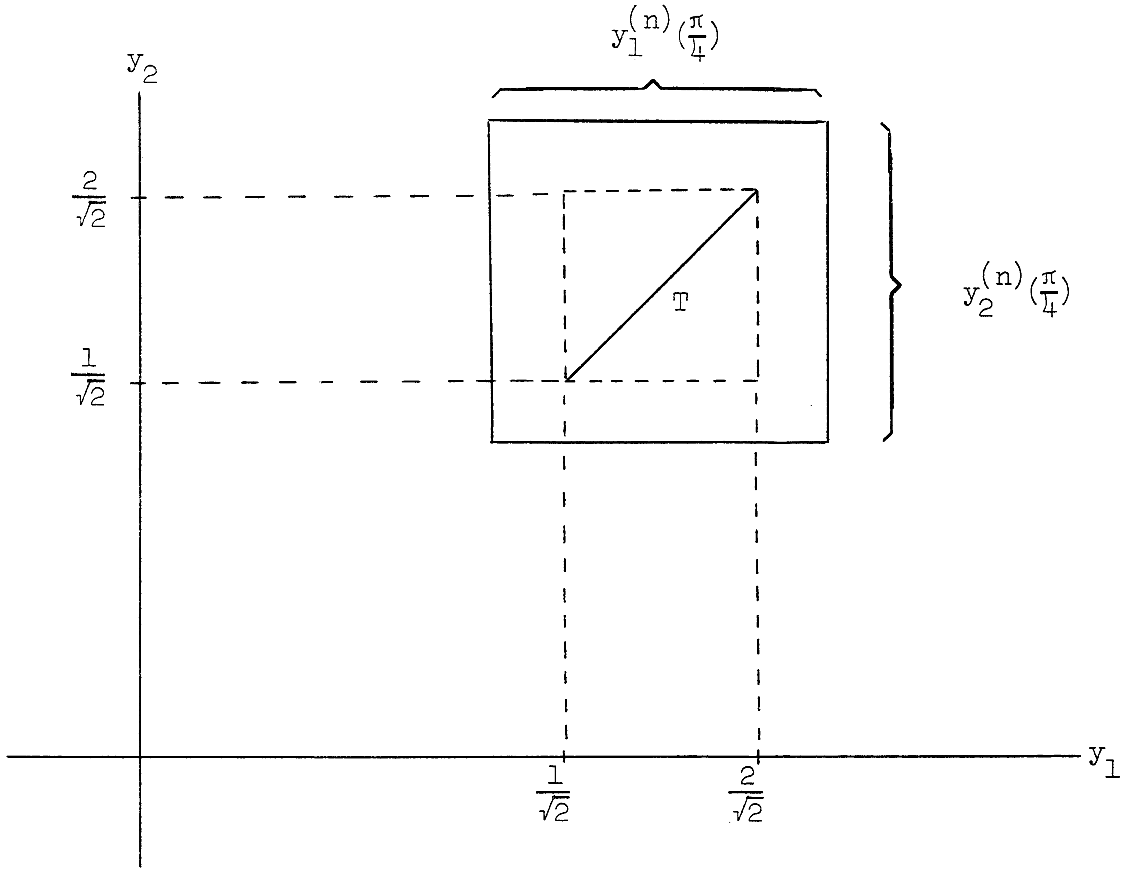

Figure 11

On the other hand, the bounding interval valued functions $y_1^{(n)}(x)$, $y_2^{(n)}(x)$ produced by the first or by the $\underline{k^{th}}$ order method yield rectangles at each x with sides parallel to the y_1 and y_2 axes which must contain the entire rectangle, $([1,2]\frac{1}{\sqrt{2}}, [1,2]\frac{2}{\sqrt{2}})$ in the y_1, y_2 plane.

The difficulty can, in principle, be dealt with by the methods we have developed. In order to describe a non-rectangular set more accurately with rectangles, we could resort to refinements. If we write

$$[1,2] = \bigcup_{r=1}^{N} \left(1 , [r - 1, r] \frac{1}{N}\right)$$

then we can find separately bounding intervals, $y_{(r)}^{(n)} (\frac{\pi}{4})$, $r = 1, 2, \ldots , N$, corresponding to the intervals of initial values $y_{20} \in 1 + [r - 1, r] \frac{1}{N}$. In this way Figure 11 is replaced by Figure 12, in which the segment T gets bounded by

$$\bigcup_{r=1}^{N} y_{(r)}^{(n)} (\frac{\pi}{r}) .$$

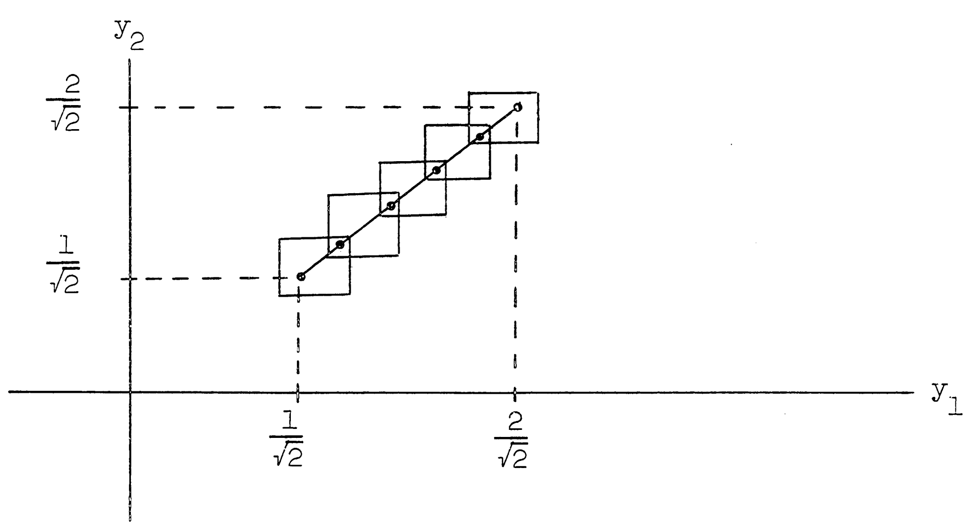

Figure 12

However, for an accurate description of a non-rectangular set in m-dimensional Euclidean space the number of small rectangles required may be very large even for reasonably small integers m.

7. Digital Computing.

In [12] a formal description of a stored program digital computer is given. The "automatic" or sequential operation of such a machine is discussed. It is shown that a computer program written for a finite*/ digital**/ interval computation will produce, as results, sets of digital intervals, each interval containing the exact result of the corresponding real arithmetic computations.

The computational procedures of the previous sections in particular can be easily modified to take into account the finite precision of digital computers. It is essentially a matter of rounding down left end points and rounding up right end points of computed intervals to the nearest digital number.

Suppose, for example, the machine arithmetic we are using in a particular program is fixed-point, signed, N-decimal-digit arithmetic with no rounding in addition or subtraction, and with rounding away from zero by one in the N^{th} place for multiplication and division if the $(N+1)^{st}$ digit of a result is five or more. More precisely, the machine product $x \otimes y$ of the signed N decimal digit numbers x, y ($|x|, |y| < 1$) is

$$x \otimes y = \operatorname{sgn}(xy) \left[|xy| \cdot 10^N + \frac{1}{2} \right] \cdot 10^{-N} ,$$

---

*/ A finite computation terminates after a finite number of "steps," i.e., elementary machine operations. (See [12], page 4.)

**/ A digital interval is one whose end points are digital numbers, i.e., are represented by machine words in some particular representation. (See [12], Section 3, and pp. 51ff.)

where $[x]$ stands for "the greatest integer less than or equal to x." (See [12], Section 4.) Also, $\operatorname{sgn}(x) = +1$ if $x > 0$, $\operatorname{sgn}(0) = 0$, and $\operatorname{sgn}(x) = -1$ if $x < 0$. If we are trying to do interval arithmetic on the computer and xy is supposed to be the right hand end point of an interval, then we will certainly have

$$xy \leq x \otimes y + 10^{-N} . \tag{7.1}$$

Hence, the digital number $x \otimes y + 10^{-N}$ will serve as right end point of an interval containing the desired interval. (See [12], Section 7.3.) Actually, in practice it is unnecessary to always round up or down; by examining the contents of the arithmetic registers and making sign tests, more accurate rounding procedures can be programmed. For example, if the $(N+1)^{st}$ digit of xy is five or more and $xy > 0$, then (7.1) can be sharpened to $xy \leq x \otimes y$.

What we obtain, then, is a "digital" or "rounded" version of interval arithmetic which can be carried out by the machine and which produces intervals $I \oplus J$, $I \ominus J$, $I \otimes J$, $I \oslash J$ with the properties

$$\begin{aligned} I + J &\subset I \oplus J \\ I - J &\subset I \ominus J \\ IJ &\subset I \otimes J \\ I/J &\subset I \oslash J . \end{aligned} \tag{7.2}$$

The quantities $I \cup J$, $I \cap J$, $w(I)$, $|I|$ can be programmed and the relations $I \subset J$, $I < J$, etc., can be verified by the machine. (See [12], Section 7.)

The digital interval arithmetic to be used on a machine can be single precision, multiple precision, or any combination of these. It can be fixed point or floating point. Again the important thing is the satisfaction of (7.2). In this way we can write machine programs for carrying out in digital interval arithmetic the computational procedures developed in the previous section of this paper. The resulting digital intervals will contain the intervals defined by those procedures and hence will contain whatever quantities lie in those intervals, since $X \in I$ and $I \subset \overline{I}$ together imply $x \in \overline{I}$.

There are finite computations with real numbers which are of interest in themselves. For example, the evaluation of real rational functions and the inversion of matrices by so-called "direct" methods (i.e., finite methods such as Gaussian elimination). For these also, corresponding computations in digital intervals contain the exact result. In this way, rigorous bounds on accumulated round-off error are obtained for any finite machine computation provided the corresponding digital interval computation can be completed. It is possible for a miscarriage to occur if a division by an interval containing zero is attempted. In this case the digital interval computation will not proceed further, even though the corresponding computation with digital real numbers might not produce a zero divisor.

In order to study in some detail the growth of interval widths due to rounding we will choose a particular form of digital arithmetic which is essentially normalized floating point binary arithmetic except that we shall ignore problems of so-called "underflow" and "overflow." In practice it is possible to program the machine to halt a computation if the exponent of a floating point result of an arithmetic operation falls outside a specified range.

Let N be a positive integer. We will assume we can represent on the computer numbers of the form

$$(7.3) \qquad (-1)^{b_o} \left(\sum_{p=1}^{N} b_p \cdot 2^{-p} \right) \cdot 2^e ,$$

where each b_p is either zero or one and e is an integer. Furthermore, either $b_1 = 1$, or else $b_o = b_1 = \cdots = b_N = e = 0$. That is, except for the number zero, every number of the form (7.3) has $b_1 = 1$. As a result the representation is unique. Let S be the set of numbers of the form (7.3). We will denote the elements of S by barred symbols when it is necessary to distinguish them from arbitrary real numbers. Since we have placed no restriction on the exponent e, it is clear that S contains all the integers n, such that $|n| < 2^N$. We define arithmetic operations in S. If

$$\bar{x} = (-1)^{b_o} \left(\sum_{p=1}^{N} b_p \cdot 2^{-p} \right) \cdot 2^e \in S$$

with $b_1 = 1$; then call

$$m(\bar{x}) = (-1)^{b_o} \sum_{p=1}^{N} b_p \cdot 2^{-p} ,$$

and $e(\overline{x}) = e$. Define $m(0) = e(0) = 0$. Thus, $m(\overline{x})$, $e(\overline{x})$ are defined for every $\overline{x} \in S$ and $\overline{x} \neq 0$ implies

$$\frac{1}{2} \leq |m(\overline{x})| < 1 .$$

For $\overline{x}, \overline{y} \in S$, consider

$$\overline{x} + \overline{y} = m(\overline{x}) \cdot 2^{e(\overline{x})} + m(\overline{y}) \cdot 2^{e(\overline{y})}.$$

Clearly, either $\overline{x} + \overline{y} = 0$, or else $\overline{x} + \overline{y}$ is uniquely representable as

$$(7.4) \qquad \overline{x} + \overline{y} = (-1)^{b_0} \left(\sum_{p=1}^{\infty} b_p \cdot 2^{-p} \right) \cdot 2^e$$

for some integer e and with each $b_p \in \{0, 1\}$ for $p = 0, 1, 2, \ldots$, such that $b_1 = 1$ and such that infinitely many of the b_p are zero. In fact, every real number is uniquely representable in this way. That is, if x is a real number, then there is a unique integer e such that $2^{e-1} \leq |x| < 2^e$, so x is of the form (7.3) for $N = \infty$, with the b_p satisfying the stated conditions. Call this integer $e(x)$ and write

$$(7.5) \qquad x = m(x)\, 2^{e(x)} .$$

Thus (7.5) defines $m(x)$ for every real number x. In particular, $m(0) = e(0) = 0$.

Finally we define a sum in S by

$$\bar{x} \oplus \bar{y} = \operatorname{sgn}(\bar{x} + \bar{y})\,[m(|\bar{x} + \bar{y}|) \cdot 2^N] \cdot 2^{-N+e(\bar{x}+\bar{y})}$$

and a difference by

$$\bar{x} \ominus \bar{y} = \bar{x} \oplus (-\bar{y}) .$$

Similarly, we define a product in S by

$$\bar{x} \otimes \bar{y} = \operatorname{sgn}(\overline{xy})\,[m(|\overline{xy}|) \cdot 2^N] \cdot 2^{-N+e(\overline{xy})}$$

and a quotient by (assuming $\bar{y} \neq 0$)

$$\bar{x} \oslash \bar{y} = \operatorname{sgn}(\overline{xy})[m(|\bar{x}/\bar{y}|) \cdot 2^N] \cdot 2^{-N+e(\bar{x}/\bar{y})} .$$

It is clear that the numbers, $\bar{x} \oplus \bar{y}$, $\bar{x} \ominus \bar{y}$, $\bar{x} \otimes \bar{y}$, $\bar{x} \oslash \bar{y}$, are again in S and can be computed by the machine. (See [12].) Furthermore, we have the inequalities

$$
\begin{aligned}
|\bar{x} \oplus \bar{y} - (\bar{x} + \bar{y})| &\leq 2^{-N+e(\bar{x}+\bar{y})} , \\
|\bar{x} \ominus \bar{y} - (\bar{x} - \bar{y})| &\leq 2^{-N+e(\bar{x}-\bar{y})} , \\
|\bar{x} \otimes \bar{y} - (\overline{xy})| &\leq 2^{-N+e(\overline{xy})} , \\
|\bar{x} \oslash \bar{y} - (\bar{x}/\bar{y})| &\leq 2^{-N+e(\bar{x}/\bar{y})} , \qquad (\bar{y} \neq 0) .
\end{aligned}
\tag{7.6}
$$

Since $x \neq 0$ implies $\frac{1}{2} \leq |m(x)| < 1$ and therefore $2^{e(x)} = \frac{x}{m(x)} \leq 2\ |x|$ we can also write

$$
(7.7) \qquad
\begin{aligned}
|\bar{x} \oplus \bar{y} - (\bar{x} + \bar{y})| &\leq 2^{-N+1}\ |\bar{x} + \bar{y}| \\
|\bar{x} \ominus \bar{y} - (\bar{x} - \bar{y})| &\leq 2^{-N+1}\ |\bar{x} - \bar{y}| \\
|\bar{x} \otimes \bar{y} - (\overline{xy})| &\leq 2^{-N+1}\ |\overline{xy}|\ , \\
|\bar{x} \oslash \bar{y} - (\bar{x}/\bar{y})| &\leq 2^{-N+1}\ |\bar{x}/\bar{y}|\ , \qquad (y \neq 0)\ .
\end{aligned}
$$

And again, since $e(\bar{x} \oplus \bar{y}) = e(\bar{x} + \bar{y})$, etc., we can write

$$
\begin{aligned}
|\bar{x} \oplus \bar{y} - (\bar{x} + \bar{y})| &\leq 2^{-N+1}\ |\bar{x} \oplus \bar{y}|\ , \\
|\bar{x} \ominus \bar{y} - (\bar{x} - \bar{y})| &\leq 2^{-N+1}\ |\bar{x} \ominus \bar{y}|\ , \\
|\bar{x} \otimes \bar{y} - (\overline{xy})| &\leq 2^{-N+1}\ |\bar{x} \otimes \bar{y}|\ , \\
|\bar{x} \oslash \bar{y} - (\bar{x}/\bar{y})| &\leq 2^{-N+1}\ |\bar{x} \oslash \bar{y}|\ , \qquad (y \neq 0)\ .
\end{aligned}
$$

Equalities occur in (7.7) and (7.8) only when both sides of an inequality are zero.

The arithmetic operations we have defined in S are not even associative. For example, let $N = 3$, $x = +(.111)\ 2^{0}$, $y = +(.101)\ 2^{-1}$, $z = +(.110)\ 2^{-2}$, then

$$
x \oplus y = +(.100)\ 2^{1} \quad \text{and} \quad (x \oplus y) \oplus z = +(.100)\ 2^{-1}\ ,
$$

whereas

$$y \oplus z = +(.100)\, 2^{0} , \qquad x \oplus (y \oplus z) = +(.101)\, 2^{1} .$$

From the above inequalities and definitions it is clear that

$$
(7.9)\qquad
\begin{aligned}
(\bar{x} \oplus \bar{y}) \ominus 2^{-N+e(\bar{x} \oplus \bar{y})} &\le \bar{x} + \bar{y} \le (\bar{x} \oplus \bar{y}) \oplus 2^{-N+e(\bar{x} \oplus \bar{y})} , \\
(\bar{x} \ominus \bar{y}) \ominus 2^{-N+e(\bar{x} \ominus \bar{y})} &\le \bar{x} - \bar{y} \le (\bar{x} \ominus \bar{y}) \oplus 2^{-N+e(\bar{x} \ominus \bar{y})} , \\
(\bar{x} \otimes \bar{y}) \ominus 2^{-N+e(\bar{x} \otimes \bar{y})} &\le \overline{xy} \le (\bar{x} \otimes \bar{y}) \oplus 2^{-N+e(\bar{x} \otimes \bar{y})} , \\
(\bar{x} \oslash \bar{y}) \ominus 2^{-N+e(\bar{x} \oslash \bar{y})} &\le \bar{x}/\bar{y} \le (\bar{x} \oslash \bar{y}) \oplus 2^{-N+3(\bar{x} \oslash \bar{y})} , \\
&\qquad\qquad (y \ne 0) .
\end{aligned}
$$

Thus, the quantities in the middle are bounded above and below by quantities which can be computed and represented by the machine.

Using the inequalities (7.9) and the formulas (2.2) and, recalling that the computer can verify the relation $\bar{x} \le \bar{y}$ and hence can compute $\max(\bar{x}, \bar{y})$, we see that digital interval operations can be programmed which not only satisfy (7.2), but also (because of (7.7))

$$
(7.10)\qquad
\begin{aligned}
w(I \oplus J) &\le w(I + J) + 2^{-N+2} |I + J| , \\
w(I \ominus J) &\le w(I - J) + 2^{-N+2} |I - J| , \\
w(I \otimes J) &\le w(IJ) + 2^{-N+2} |IJ| , \\
w(I \oslash J) &\le w(I/J) + 2^{-N+2} |I/J| , \qquad (0 \notin J) .
\end{aligned}
$$

From the inequalities (7.10) and Theorem 4.1 it is clear that for any finite digital interval computation corresponding to the evaluation of a rational interval function F with real valued real restriction f that an interval I_N is produced whose width for any given point in D_f which is digitally representable is bounded by $K \cdot 2^{-N}$ for some number K. Therefore, the width of I_N can be made arbitrarily small for large enough N. Put in other words, if we start with a set of intervals of zero width and perform some finite sequence of digital interval arithmetic operations beginning with the given set of intervals, then the digital intervals obtained will have widths of order 2^{-N}. That is, if the same computation is performed with various values of N, then the width of each resulting interval will be bounded by some constant independent of N times 2^{-N}.

It also follows from (7.10) and Theorem 4.1, more generally, that the excess in width of a digitally computed interval over the width of an interval resulting from a finite interval arithmetic computation using exact real arithmetic to obtain end points is also of the order 2^{-N}.

The significance of digital interval arithmetic is further clarified by its application to some particular computational schemes.

The so-called Gaussian elimination procedure for solving a system of linear algebraic equations

$$\sum_{j=1}^{n} a_{ij} x_j = b_i , \qquad i = 1, 2, \ldots, n ,$$

consists first in transforming the given system to an equivalent one in which the matrix of coefficients of the x_j is made zero below the main diagonal. This is done by forming linear combinations of the given equations. Then the resulting system can be solved directly for the x_j, starting with x_n and regressing to x_1.

The procedure is described formally as follows: Define

$$a_{ij}^{(1)} = a_{ij} , \qquad b_i^{(1)} = b_i \qquad (i, j = 1, \ldots , n) .$$

For $p = 1, 2, \ldots , n-1$, define

$$\left.\begin{aligned} a_{ij}^{(p+1)} &= a_{ij}^{(p)} - \left(\frac{a_{ip}^{(p)}}{a_{pp}^{(p)}}\right) a_{pj}^{(p)} \\ b_i^{(p+1)} &= b_i^{(p)} - \left(\frac{a_{ip}^{(p)}}{a_{pp}^{(p)}}\right) b_p^{(p)} \end{aligned}\right\} \quad n \geq i , \quad j \geq p+1 .$$

Then

$$x_n = b_n^{(n)} / a_{n,n}^{(n)}$$

and

$$x_p = \left(b_p^{(p)} - \sum_{j=p+1}^{n} a_{pj}^{(p)} x_j\right) \bigg/ a_{pp}^{(p)}$$

for $p = n-1, n-2, \ldots , 2, 1$. By reordering the components of the solution vector it can be arranged that

$$|a_{pp}^{(p)}| \geq |a_{pj}^{(p)}| ,$$

for $j \geq p$. Should $a_{pp}^{(p)} = 0$ for some p, then the given matrix has zero determinant and the procedure fails.

Given intervals A_{ij}, B_i $(1 \leq i, j \leq n)$ we can compute X_n, X_{n-1}, $\ldots$, X_2, X_1, as defined by the procedure above interpreting the arithmetic operations as interval arithmetic operations. Providing the intervals $A_{pp}^{(p)}$ so computed do not contain the real number zero for $p = 1, 2, \ldots, n$, this scheme will produce intervals X_j, $j = 1, 2, \ldots, n$, such that

$$\sum_{j=1}^{n} a_{ij} x_j = b_i ,$$

with $a_{ij} \in A_{ij}$, $b_i \in B_i$, $i \leq i, j \leq n$ implies $x_j \in X_j$, $j = 1, 2, \ldots, n$.

For machine computation, we replace exact arithmetic operations by their corresponding digital (or rounded) versions and the digital interval version of the procedure which results is described by:

$$\overline{A}_{ij}^{(1)} \supset A_{ij} , \qquad \overline{B}_i^{(1)} \supset B_i , \qquad (i, j = 1, 2, \ldots, n)$$

(the initial coefficients A_{ij}, B_i may not be exactly representable digitally, e.g., $A_{11} = [1/3, \pi]$) and for $p = 1, 2, \ldots, n-1$

$$\overline{A}_{ij}^{(p+1)} = \overline{A}_{ij}^{(p)} \ominus \left\{ \left(\overline{A}_{ip}^{(p)} \oplus\!\!\!\!\div\; \overline{A}_{pp}^{(p)} \right) \otimes \overline{A}_{pj}^{(p)} \right\} ,$$

$$\overline{B}_i^{(p+1)} = \overline{B}_i^{(p)} \ominus \left\{ \left(\overline{A}_{ip}^{(p)} \oplus\!\!\!\!\div\; \overline{A}_{pp}^{(p)} \right) \otimes \overline{B}_p^{(p)} \right\} .$$

Then

$$\overline{X}_n = \overline{B}_n^{(n)} \oplus\!\!\!\!\div\; \overline{A}_{n,n}^{(n)} ,$$

and for $p = n-1, n-2, \ldots, 2, 1,$

$$\overline{X}_p = \left\{ \overline{B}_p^{(p)} \ominus \sum_{j=p+1}^{n} \overline{A}_{pj}^{(p)} \otimes \overline{X}_j \right\} \oslash \overline{A}_{pp}^{(p)} ,$$

where $\sum$ stands for digital interval summation. The evaluation of (7.11) affords an opportunity to explain the use of mixed precision computations with digital intervals. Just as is common practice in real digital computing, we can sum or "accumulate" the products, $\overline{A}_{pj}^{(p)} \otimes \overline{X}_j$ using double precision (digital interval) arithmetic with a resulting reduction in the width of the interval sum. If $0 \in \overline{A}_{pp}^{(p)}$ for some $1 \leq p \leq n$, then the machine will halt, unable to decide whether a zero determinant can occur for some set of $a_{ij} \in \overline{A}_{ij}$. Otherwise, i.e., if $0 \notin \overline{A}_{pp}^{(p)}$ for $p = 1, 2, \ldots, n$, the procedure yields intervals $\overline{X}_j$ with $x_j \in \overline{X}_j$ for every set of x_j, $j = 1, 2, \ldots, n$, such that

$$\sum_{j=1}^{n} a_{ij} x_j = b_i \qquad (i \leq i, \quad j \leq n)$$

for some set of a_{ij}, b_i with $a_{ij} \in \overline{A}_{ij}$, $b_i \in \overline{B}_i$.

If $w(\overline{A}_{ij}) = w(\overline{B}_i) = 0$, $1 \leq i, j \leq n$, that is, if the initial coefficients are all intervals of zero width, and if $0 \notin \overline{A}_{pp}^{(p)}$, $p = 1, 2, \ldots, n$, then we know that there is a number K such that $w(\overline{X}_j) \leq K\, 2^{-N}$ $(j = 1, 2, \ldots, n)$. Therefore, with high enough precision, i.e., for large enough N, the digital interval procedure will produce intervals $\overline{X}_j$ of arbitrarily small width. Notice that $0 \notin \overline{A}_{pp}^{(p)}$ for a particular value of N implies $0 \notin \overline{A}_{pp}^{(p)}$ for all larger N, by inclusion monotonicity of interval arithmetic and by virtue of the fact that $|x \otimes y - xy|$, $|x \oslash y - x/y|$, etc., are monotonic decreasing as N increases.

Actual machine computations were made inverting matrices using digital interval arithmetic and an elimination procedure similar to the above described one. The binary equivalent of about 8 decimal places was the machine precision used $(N = 27)$. Some 14 x 14 matrices were inverted with resulting intervals containing the coefficients of the inverse matrix of relative width about 10^{-1}. This was a loss of about seven decimal places, so to speak. The exact values of the coefficients of the inverse matrix were known and it turned out that the mid points of the computed intervals were much closer to the exact results than indicated by their width. It is not clear at present whether this was an "accident" or whether there is a tendency for this particular digital interval matrix inversion procedure to grow intervals fairly symmetrically about the infinite precision result. There is some reason to suspect the latter since interval subtractions and divisions reverse the roles of the end points of a given interval. (See (2.2) and also pp. 112ff below.)

Another comment on the numerical results quoted is in order. Von Neumann and Goldstine ([15], pp. 1023), speaking of a direct method which is a slight variant of Gaussian elimination (assuming, however, fixed point arithmetic rather than the floating point arithmetic which was used in the numerical work quoted), summarize the result of their round-off analysis with the statement, "Matrices of order 15, 50, 150 can usually be inverted with a (relative) precision of 8, 10, 12 decimal digits less respectively than the number of digits carried throughout."

Since most computers employ 8 or 9 decimal digit relative precision in standard single precision floating point operation, they should be able to invert "most" matrices with satisfactorily small interval width using double precision digital interval arithmetic in connection with a direct elimination procedure such as described above. And in any case, the computed intervals, narrow or wide, always contain the exact solution to the given linear system.

We have already seen (pages 8 and 13 above) that rational interval functions can have as values intervals whose widths exceed the widths of the corresponding values of the united extension of their real restriction. And we have seen (pages 21 and 22 above) that the excess has something to do with the multiple occurrence of one or more variables in the expression for the rational interval function.

The technique of refinements allows the diminution of excess interval width. However, it is clear that it would be hopelessly inefficient to make extensive use of this device in practice in connection with digital computations involving a very large number of independent variables — for example, in matrix inversion. The device will be useful in digital computations only when there are few variables which have a large number of occurrences.

We will illustrate these remarks by considering some computations involving "contraction" mappings.

Suppose F is a rational interval function with domain $\mathcal{I}_A$ and that $\overline{F}$ is a digital majorant with domain $\mathcal{I}_{\overline{A}}$, $\overline{A} \subset A$; that is, $\overline{F}(\overline{X})$ is expressible in digital interval arithmetic as a function of $\overline{X}$ with $X \subset \overline{X}$ implying $F(X) \subset \overline{F}(\overline{X})$. If there is a real number $k < 1$, such that for $X \subset A$ we have $F(X) \subset F(A)$ and

$$w(F(X)) \leq k\, w(X) ,$$

then F is said to be a "strong contraction" mapping and from $X \subset A$ it follows that $F(X) \subset F(A)$, by inclusion monotonicity of F. Define $G(X_o, 0) = X_o$ and, for $n \geq 1$, define $G(X_o, n) = F(G(X_o, n-1))$. Then

$$w(G(X_o, n)) \leq k^n\, w(X_o) ,$$

so that for every $X_o \subset A$ we have $G(X_o, n) \subset G(A, n)$, and there is a real number $y \in F(A)$ such that for every $X_o \subset A$,

$$\lim_{n \to \infty} G(X_o, n) = y = \lim_{n \to \infty} G(A, n) .$$

The number y satisfies $y = F(y) \in G(A, n)$. By previous discussion, the digital majorant $\overline{F}$ will satisfy

$$w(\overline{F}(\overline{X})) \leq k\, w(\overline{X}) + k' \, 2^{-N}$$

for some k', which is independent of $\overline{X}$ (by virtue of the compactness of $\mathcal{J}_{\overline{A}}$). Call $\overline{G}(\overline{X}_o, 0) = \overline{X}_o$ and $\overline{G}(\overline{X}_o, n) = \overline{F}(\overline{G}(\overline{X}_o, n-1))$. Then

<u>Theorem</u> 7.1:

$$w(\overline{G}(\overline{X}_o, n)) \leq k^n\, w(\overline{X}_o) + \frac{1 - k^n}{1 - k}\, k' \cdot 2^{-N} .$$

Now $\overline{F}$ is also inclusion monotonic, so if $\overline{F}(\overline{A}) \subset \overline{A}$, then there is a digital interval $\overline{Y} \subset \overline{F}(\overline{A})$ such that for $\overline{X}_o \subset \overline{A}$, we have

$$\lim_{n \to \infty} \overline{G}(\overline{X}_o, n) = \overline{Y} = \lim_{n \to \infty} \overline{G}(\overline{A}, n)$$

with $w(\overline{Y}) \leq \frac{k'}{1 - k} \cdot 2^{-N}$. Notice also that for each n we have

$$\overline{Y} \subset \overline{G}(\overline{A}, n) .$$

As an example, consider the rational interval function F on $\mathcal{I}_{[1,2]}$ defined by

$$F(X) = 1 + \frac{1}{1 + X} .$$

We have

$$\begin{aligned} F([1,2]) &= 1 + \frac{1}{1 + [1,2]} \\ &= 1 + \frac{1}{[2,3]} \\ &= 1 + \left[\frac{1}{3} , \frac{1}{2}\right] \\ &= \left[\frac{4}{3} , \frac{3}{2}\right] \subset [1,2] , \end{aligned}$$

and for $X \subset [1,2]$, we have

$$w(F(X)) \leq \left|\frac{1}{1 + X}\right|^2 w(X) \leq \frac{1}{4} w(X) .$$

So we find in particular, that for

$$X_o = [1,2] , \qquad X_{n+1} = F(X_n) ,$$

we have

$$w(X_n) \leq (\tfrac{1}{4})^n ,$$

and so the intervals X_n converge to, and each of them contains, the fixed point of f on $[1,2]$ which is the number $\sqrt{2}$, since $x = 1 + \frac{1}{1+x}$ implies $(x - 1)(1 + x) = 1$ or $x^2 = 2$.

Using 3 place decimal digit arithmetic for the digital interval version it can be shown that

$$\overline{X}_o = [1,2] , \qquad \overline{X}_{n+1} = 1 \oplus \{1 \oslash (1 \oplus \overline{X}_n)\}$$

implies that the intervals $\overline{X}_n$ converge to $[1.41, 1.42]$. In fact, $\overline{X}_n = [1.41, 1.42]$ for $n \geq 3$. We show this by direct computation. For $n = 1$, we have

$$\overline{X}_1 = 1 \oplus \{1 \oslash (1 \oplus [1,2])\}$$

$$= 1 \oplus \{1 \oslash [2,3]\}$$

$$= 1 \oplus [.333, .500] = [1.33, 1.50] ;$$

for $n = 2$, we have

$$\overline{X}_2 = 1 \oplus \{1 \oslash (1 \oplus [1.33, 1.50])\}$$

$$= 1 \oplus \{1 \oslash [2.33, 2.50]\}$$

$$= 1 \oplus [.400, .430] = [1.40, 1.43] ;$$

for $n = 3$, we have

$$\overline{X}_3 = 1 \oplus \{1 \oslash (1 \oplus [1.40, 1.43])\}$$

$$= 1 \oplus \{1 \oslash [2.40, 2.43]\}$$

$$= 1 \oplus [.411, .417] = [1.41, 1.42] ;$$

and for $n \geq 3$, we have $\overline{X}_n = [1.41, 1.42]$, since

$$\overline{X}_n = 1 \oplus \{1 \oslash (1 \oplus [1.41, 1.42])\}$$

$$= 1 \oplus \{1 \oslash [2.41, 2.42]\}$$

$$= 1 \oplus [.413, .415] = [1.41, 1.42] .$$

The real restriction of the function F in the above example is the real function f given by $f(x) = 1 + \frac{1}{1 + x}$ with $x \in [1,2]$. Notice that f is also a strong contraction, i.e., f maps the interval $[1,2]$ into itself and

$$|f(x_1) - f(x_2)| \leq \frac{1}{4} |x_1 - x_2|$$

for $x_1, x_2 \in [1,2]$.

On the other hand, the real rational function g as defined on $[1,2]$ by

$$g(x) = \frac{x}{2} + \frac{1}{x}$$

is a strong contraction since g maps $[1,2]$ into itself and

$$|g(x_1) - g(x_2)| \leq \frac{1}{2} |x_1 - x_2| ,$$

whereas the interval extension G as defined on $\mathcal{I}_{[1,2]}$ by

$$G(X) = (X/2) + \frac{1}{X} ,$$

is not a strong contraction. While $G([1,2]) = [1,2] \subset [1,2]$, it happens that $w(G(X)) \leq k\, w(X)$ for all $X \subset [1,2]$ implies $k \geq 1$. In fact, if $X = [x_1, x_2] \subset [1,2]$, then

$$\begin{aligned} G(X) &= \frac{[x_1 , x_2]}{2} + \frac{1}{[x_1 , x_2]} \\ &= \left[\frac{x_1}{2} + \frac{1}{x_2} , \frac{x_2}{2} + \frac{1}{x_1} \right] \end{aligned}$$

(notice the mixing of end points), and

$$(7.12) \quad w(G(X)) = (x_2 - x_1) \left(\frac{1}{2} + \frac{1}{x_1 x_2} \right) = \left(\frac{1}{2} + \frac{1}{x_1 x_2} \right) w(X) .$$

In particular, for $x_1 = 1$, any X of the form $X = [1, x_2]$ with $x_2 \in [1,2]$ fails to satisfy $w(G(X)) < w(X)$.

If $\overline{g}$ is the united extension of g, that is, if

$$\overline{g}(X) = \left\{ \frac{x}{2} + \frac{1}{x} \mid x \in X \right\} = \left\{ g(x) \mid x \in X \right\}$$

then

$$w(\overline{g}(X)) = \max_{x_1, x_2 \in X} |g(x_1) - g(x_2)| \leq \frac{1}{2} w(X) .$$

The technique of refinements allows us to approximate $\overline{g}$ by refinements of G. For $X \subset [1,2]$ with $X = [x_1, x_2]$, define

$$X_i^{(n)} = x_1 + [i - 1, i] \frac{x_2 - x_1}{n} ,$$

then the refinements $G^{(n)}$ of G are defined by

$$G^{(n)}(X) = \bigcup_{i=1}^{n} G(X_i^{(n)}) .$$

For each n, $G^{(n)}$ is an interval values function on $\mathcal{J}_{[1,2]}$ also having g as its real restriction. In fact (see Section 4 above),

$$\overline{g}(X) \subset G^{(n)}(X)$$

for all $X \subset [1,2]$. Now $G^{(n)}(X) \subset G(X)$, so

$$G^{(n)}([1,2]) \subset [1,2] = G([1,2]) .$$

Furthermore,

$$w(G^{(n)}(X)) \leq \sum_{i=1}^{n} w(G(X_i^{(n)})) .$$

By (7.12), we have

$$w(G(X_i^{(n)})) \leq \left(\frac{1}{2} + \left(\frac{1}{X_i^{(n)}}\right)^2\right) \frac{w(X)}{n} .$$

Now

$$\sum_{i=1}^{n} \left|\frac{1}{X_i^{(n)}}\right|^2 < \int_{X - \frac{w(X)}{n}} \frac{1}{x^2}\, dx < \frac{w(X)}{\left|X - \frac{w(X)}{n}\right|^2} ,$$

so

$$\sum_{i=1}^{n} \left|\frac{1}{X_i^{(n)}}\right|^2 < \frac{1}{(1 - \frac{1}{n})^2} ,$$

and therefore

$$w(G^{(n)}(X)) < \frac{w(X)}{n}\left(\frac{n}{2} + \frac{1}{(1 - \frac{1}{n})^2}\right)$$

$$< w(X)\left(\frac{1}{2} + \frac{1}{n(1 - \frac{1}{n})^2}\right) .$$

Now

$$\frac{1}{n(1 - \frac{1}{n})^2} = \frac{1}{n - 2 + \frac{1}{n}} < \frac{1}{n - 2} ,$$

so if $n \geq 4$, then

$$w(G^{(n)}(X)) \le k_n\, w(X) ,$$

with

$$k_n = \frac{1}{2} + \frac{1}{n(1 - \frac{1}{n})^2} < 1.$$

Therefore, the refinements $G^{(n)}$ of G are strong contractions for $n \ge 4$ and, in particular, we can iterate say, $G^{(4)}$ to obtain

$$X_1 = [1,2] , \qquad X_{k+1} = G^{(4)}(X_k) , \qquad k = 1 , 2 , \ldots ,$$

with the result that for each $k = 1, 2, \ldots$ we have $\sqrt{2} \in X_k$ and $w(X_k) \to 0$. Actually, convergence of $w(X_k)$ to zero is also obtained for $n = 2$ and $n = 3$.

The failure of $w(X_k)$ to converge to zero for $n = 1$ cannot be remedied by starting with a smaller X_o containing $\sqrt{2}$. For we find that, choosing $\epsilon_1, \epsilon_2 > 0$ (according to (7.12)), we have for $X = \sqrt{2} + [-\epsilon_1, \epsilon_2]$

$$w(G(X)) = \left(\frac{1}{2} + \frac{1}{(\sqrt{2} - \epsilon_1)(\sqrt{2} + \epsilon_2)}\right) w(X) ,$$

so that, for every small $\epsilon_1, \epsilon_2,$ we have

$$w(G(X)) = \left(1 + \frac{1}{2\sqrt{2}} (\epsilon_1 - \epsilon_2) + \cdots\right) w(X) ,$$

and $w(G(X)) < w(X)$ only if $\epsilon_1 < \epsilon_2$. However, for very small $\epsilon_1, \epsilon_2,$ we also have

$$G(\sqrt{2} + [-\epsilon_1, \epsilon_2]) = \sqrt{2} + \left[-\frac{(\epsilon_1 + \epsilon_2)}{2} + \cdots , \frac{(\epsilon_1 + \epsilon_2)}{2} + \cdots \right]$$

and $\epsilon_1 < \epsilon_2$ implies

$$-\frac{(\epsilon_1 + \epsilon_2)}{2} < -\epsilon_1 ,$$

so

$$G(\sqrt{2} + [-\epsilon_1, \epsilon_2]) \not\subset \sqrt{2} + [-\epsilon_1, \epsilon_2] .$$

For a numerical example of the digital version of iterating a refinement, we will choose N decimal place floating point arithmetic again and define, for $\overline{X} \subset [1,2]$,

$$\overline{G}(\overline{X}) = (\overline{X} \oplus\!\!\!\!\div\, 2) \oplus (1 \oplus\!\!\!\!\div\, \overline{X}) .$$

If $\overline{X} = [\overline{x}_1, \overline{x}_2]$, then we <u>could</u> define $\overline{X}_i^{(n)}$ by

$$\overline{X}_i^{(n)} = \overline{x}_1 \oplus \left\{ [i - 1, i] \otimes \left((\overline{x}_2 \ominus \overline{x}_1) \oplus\!\!\!\!\div\, n \right) \right\} ,$$

treating even the digital numbers $\overline{x}_1, \overline{x}_2, n$ as digital intervals. We would have $\overline{X}_i^{(n)} \supset X_i^{(n)}$. But a better way of constructing $\overline{X}_i^{(n)}$ is to prevent overlap. To this end, we define

$$\overline{X}_i^{(n)} = \left[\overline{x}_i^{(n)} , \overline{y}_i^{(n)} \right]$$

with

$$\bar{x}_1^{(n)} = \bar{x}_1 , \quad \bar{x}_{i+1}^{(n)} = \bar{y}_i^{(n)} , \qquad i = 1 , 2 , \ldots , n-1 ,$$

and

$$\bar{y}_n^{(n)} = \bar{x}_2 , \quad \bar{y}_i^{(n)} = |\bar{x}_i^{(n)} \oplus \{(\bar{x}_2 \ominus \bar{x}_1) \oslash n)\}|$$

for $i = 1, 2, \ldots , n-1$. We assume that the number of digits N in the digital arithmetic used is large enough or, equivalently, that n is small enough so that

$$\bar{y}_{n-1}^{(n)} < \bar{y}_n^{(n)} .$$

For example, if $n = 2$, and $N = 3$, then we compute as follows:

Let $\bar{X}_o = [1,2]$, and define

$$\bar{X}_{k+1} = \bar{G}\left((\bar{X}_k)_1^{(2)}\right) \bigcup G\left((\bar{X}_k)_2^{(2)}\right) .$$

We find that $|1 \oplus \{(2 \ominus 1) \oslash 2\}| = 1.5$, so

$$(\bar{X}_o)_1^{(2)} = [1, 1.5] , \quad (\bar{X}_o)_1^{(2)} = [1.5, 2] ,$$

and

$$\bar{G}([1, 1.5]) = [1.16, 1.75] , \quad \bar{G}([1.5, 2]) = [1.25, 1.67] ,$$

so

$$\overline{X}_1 = [1.16, 1.75] .$$

Then

$$|1.16 \oplus \{(1.75 \ominus 1.16) \oslash 2\}| = 1.46 ,$$

so

$$(\overline{X}_1)_1^{(2)} = [1.16, 1.46] , \qquad (\overline{X}_1)_2^{(2)} = [1.46, 1.75] ,$$

and

$$\overline{G}([1.16, 1.46]) = [1.26, 1.60] , \quad \overline{G}([1.46, 1.75]) = [1.30, 1.56] ,$$

so

$$\overline{X}_2 = [1.26, 1.60] .$$

Similarly, we find that

$$\overline{X}_3 = [1.32, 1.51] , \qquad \overline{X}_4 = [1.36, 1.47] ,$$

$$\overline{X}_5 = [1.38, 1.45] , \qquad \overline{X}_6 = [1.39, 1.44] ,$$

$$\overline{X}_7 = [1.39, 1.43] , \qquad \overline{X}_8 = [1.40, 1.43] ,$$

$$\overline{X}_9 = \overline{X}_{10} = \cdots = [1.40, 1.43] .$$

Compare these computations with those on pages 109-118 and with those on page 12. We have three interval arithmetic computational schemes producing convergent sequences of intervals all containing the number, $\sqrt{2}$. (In the case of the digital arithmetic computations the convergence is to an interval of positive width.) These three schemes appear very similar but turn out to be quite different with respect to efficiency.

At this point it should already be clear that machine programs can be written for generating digital intervals containing the range of values for the commonly used functions such as $\sqrt{x}$, $\exp x$, $\ell n\ x$, $\sin x$, etc., when the argument varies over a digital interval $\overline{X}$. This can be done efficiently using interval computations based on quickly convergent real approximations. The excess widths of the resulting intervals can be held down using special properties of the functions concerned. For example, $|\sin x| < 1$, $\sin(x + 2\pi) = \sin x$, $\sqrt{x}$ is monotonic, etc. And, of course, by resorting to multiple precision the excess width can be made arbitrarily small.

In Section 5 above we discussed two families of k^{th} order interval methods for computing definite integrals. One was based on local expansion of the integrand in Taylor's series and the other was the Gaussian procedure. In both cases the interval method was made possible by the appearance of the real approximation method as a finite sum of functions derivable in an appropriate way from the integrand.

In Section 6 we presented but a single family of k^{th} order interval methods, $k = 1, 2, 3, \ldots$, for the solution of systems of ordinary differential equations of first order. The methods were based on local expansion in Taylor's series and required in addition to the given functions

$$f_j(x, y_1, \ldots, y_m) = \frac{dy_j}{dx},$$

the functions $f_j^{(\ell)}$, $\ell = 1, 2, \ldots, k-1$, for the k^{th} order method. Recall that

$$f_j^{(\ell)} = \frac{d}{dx} f_j^{(\ell-1)} .$$

The possibility of using local Taylor expansion to approximate the solutions to ordinary differential equations with ordinary real computations has been noticed already, of course. It has also been noticed ([8], page 66) that the expressions for the functions $f_j^{(\ell)}$ increase with ℓ in complexity. Actually, for <u>rational</u> f_j, the computer can be programmed to generate <u>values</u> of $f_j^{(\ell)}$ by use of recursion formulas so that one need only supply the functions f_j, themselves. For example, if

$$y' = f(x, y) = x^2 + y^2 ,$$

then a value of

$$f^{(\ell)}(x, y) , \qquad \ell \geq 1 ,$$

can be generated from the numbers x, y, $f(x, y)$ by computing successively

$$f^{(1)}(x, y) = 2(x + y\, f(x, y)) ,$$

$$f^{(2)}(x, y) = 2(1 + y f^{(1)}(x, y) + (f(x, y))^2$$

$$\cdots$$

$$f^{(\ell+1)}(x, y) = 2 \left\{ \sum_{r=0}^{\ell} \binom{\ell}{r} f^{(r-1)}(x, y) f^{(\ell-r)}(x, y) \right\} , \qquad \ell > 1 ,$$

with $f^{(-1)}(x, y)$ defined by $f^{(-1)}(x, y) = y$,

and $f^{(0)} (x, y)$ defined by $f^{(0)} (x, y) = f(x, y)$.

One need never actually write down an expression for $f^{(\ell)}(x, y)$ in terms of x and y. In fact, a general program can be written for generating values of functions

$$f_j^{(\ell)} = \frac{d^\ell}{dx^\ell} f_j ,$$

from the coefficients of the rational expressions for $f_j(x, y_1, \ldots , y_m)$, $j = 1, 2, \ldots , m$, and the numbers $x, y_1, \ldots , y_m$, $f_1(x, y_1, \ldots , y_m)$, $\ldots , f_m(x, y_1, \ldots , y_m)$.

Instead of using Taylor expansions one could develop interval methods of k^{th} order which bound the solutions to ordinary differential equations based on almost any real finite sum approximation method with a suitable expression for the local truncation error of order k (in powers of the stepwidth).

Euler's method, for example approximates

$$y(x + h) = y(x) + h f(x, y(x)) + \frac{1}{2} h^2 f^{(1)}(t, y(t)) ,$$

with

$$t \in [x, x + h]$$

by

$$y_{i+1} = y_i + h\, f(x_i, y_i)$$

with a local truncation error of order 2 (in powers of h), namely

$$T_i = \frac{1}{2} h^2 f^{(1)}(t, y(t)) ,$$

for some

$$t \in [x_i, x_i + h] .$$

The general process we have in mind produces out of this particular method exactly what we obtained as our 2^{nd} order method based on Taylor expansion, i.e., read (6.13) and (6.14) with $k = 2$. We see from this that the extra work paid out in evaluations of the local truncation expression in interval arithmetic (over the rectangular boxes constructed as part of the interval method; see Figure 9, page 68) buys bounding intervals whose widths are of the same order in powers of the step size as the local truncation error of the real method upon which the interval method is based. In this sense, our first order interval method*/ can

*/ See page 61.

be viewed as an interval method based on the ridiculously simple real method defined by $y_{i+1} = y_i$ with a local truncation error expressed by $T_i = h\, f(t, y(t))$ for some $t \in [x_i, x_i + h]$.

Of course, not all interval methods so derived from real methods will turn out identical with some member of our family of "Taylor made" methods. For example, the local truncation error for the well known Runge-Kutta method is not a total derivative of the function f, but rather a certain combination of partial derivatives of f. (See e.g., [8], pp. 127-132.) Furthermore, the approximation formula involves a sum of values of f alone; to wit,

$$y_{i+1} = y_i + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) ,$$

where

$$k_1 = f(x_i, y_i) ,$$

$$k_2 = f(x_i + \frac{h}{2} , y_i + \frac{h}{2} k_1)$$

$$k_3 = f(x_i + \frac{h}{2} , y_i + \frac{h}{2} k_2)$$

$$k_4 = f(x_i + h , y_i + h\, k_3) .$$

This, as many other real methods do, has the form

$$Y(x) = a(x, x_i, y_i, x_{i-1}, y_{i-1}, \ldots , x_{i-s}, y_{i-s}) ,$$

with $x = x_i + h$ for some fixed non-negative integer s, with a local truncation error of the form $T(t_i, u_i)\, h^k$ for some integer k with

$$t_i = t(x, x_i, y_i, \ldots, x_{i-s}, y_{i-s}) ,$$

$$u_i = u(x, x_i, y_i, \ldots, x_{i-s}, y_{i-s}) ,$$

for some $(t_i, u_i) \in D_f$.

We will outline here a possible approach for these methods. Proceeding as in Section 6, we choose a rectangular domain, $[x_o, a] \otimes B_1 \otimes \cdots \otimes B_m$, called D_f in which the functions f_j (occurring in (6.1)) and also the functions T_j (the components of T) are continuous. Assume interval extensions $\tilde{a}$ and $\tilde{T}$ of a and T can be found which satisfy conditions 1), 2), and 3) of page 59 above. Call $p_i = (x_i, y_i)$ and $q_i = (t_i, u_i)$, then replace (6.13) by

$$A(x, p_i, \ldots, p_{i-s}) = \tilde{a}(x, p_i, \ldots, p_{i-s}) + \tilde{T}(D_f)(x - x_i)^k$$

and (6.14) by

$$y_i^{(n)}(x) = A(x, p_o, \ldots, p_{-s}) \bigcap \left\{ \tilde{a}(x, p_{i-1}, \ldots, p_{i-s-1}) + \tilde{T}\left(X_i^{(n)}, A(X_i^{(n)}, p_{i-1}, \ldots, p_{i-s-1})\right)(x - x_{i-1})^k \right\}$$

We turn finally to the discussion of the digital version of the family of k^{th} order methods discussed in Section 6 above (essentially defined by (6.13) and (6.14)). We assume digital interval arithmetic operations $\oplus$, $\ominus$, $\otimes$, $\oslash$, as described earlier in this section.

It is sufficient for our purpose here to consider the special case of a single differential equation of first order,

$$y' = f(x, y) ,$$

with initial condition $y(0) = y_0$; i.e., $x_0 = 0$. Suppose now that F, $F^{(1)}, \dots , F^{(k-1)}$, satisfy the conditions of pages 71 and 72. In particular,

$$w(F^{(\ell)}(X, Y)) \leq K^{(\ell)} \max (w(X), w(Y)) ,$$

with $K^{(\ell)} > 0$ for $\ell = 0, 1, 2, \dots , k-1$. Recall that $F^{(\ell)}(x, y) = f^{(\ell)}(x, y)$ for $(x, y) \in D_f$. Let $\overline{F}^{(\ell)}$ be digital interval majorants to $F^{(\ell)}$, $\ell = 0, 1, \dots , k-1$, such that $F^{(\ell)}(\overline{X}, \overline{Y}) \subset \overline{F}^{(\ell)}(\overline{X}, \overline{Y})$ and

$$w(\overline{F}^{(\ell)}(\overline{X}, \overline{Y})) \leq w(F^{(\ell)}(\overline{X}, \overline{Y})) + \overline{K}^{(\ell)} 2^{-N} ,$$

for some $\overline{K}^{(\ell)}$ independent of $\overline{X}, \overline{Y}$ in $D_{\overline{F}} \subset D_F$. This is possible, for example, for rational F since $D_{\overline{F}}$ is compact.

The interval polynomials in (6.13) and (6.14) could just as well have been defined in nested form (see Section 2). We chose in Section 6 not to do this for reasons of conceptual simplicity. It is clear that the results of Section 6 are essentially unchanged with this modification. For computation purposes, the nested evaluation of interval polynomials produces intervals of smaller width generally than the sum of powers evaluation (see Section 2).

Define $\overline{A}$ on $\mathcal{J}_{[o,a]} \otimes D_{\overline{F}}$ by finite induction as follows:

$$\overline{P}_o = \overline{F}^{(k-1)}(\overline{D}_f) ,$$

and for $j = 0, 1, \ldots , k-1,$

$$\overline{P}_{j+1} = \left\{ (\overline{P}_j \oplus (k - j)) \otimes (\overline{X} \ominus \overline{x}) \right\} \oplus \overline{F}^{(k-2-j)}(\overline{x}, \overline{Y}) ,$$

where

$$\overline{F}^{(-1)}(\overline{x}, \overline{Y}) = \overline{Y} .$$

Then

$$\overline{A}(\overline{X}, \overline{x}, \overline{Y}) = \overline{P}_k .$$

If $D_{\overline{F}} = \mathcal{J}_{[o,\overline{a}]} \otimes \mathcal{J}_{\overline{B}}$, then $\overline{D}_f = [0, \overline{a}] \times \overline{B}$. Using $\overline{A}$ in place of A and $\overline{B}$ in place of B, suppose $\overline{X}^*$ is determined by the process described in Section 6, so that $\overline{X}^* \subset [0, \overline{a}]$ and $\overline{A}(\overline{X}^*, 0, \overline{y}_o) \subset \overline{B}$, with $y_o \in \overline{y}_o$. Call $\overline{w} = w(\overline{X}^*)$.

The digital arithmetic we are using is binary based. We will make the simplifying assumption that the integer n, denoting the number of slices to be made out of $\overline{X}^*$, is of the form 2^s for a positive integer s, and that the intervals $X_i^{(n)}$, defined by

$$X_i^{(n)} = [i - 1, i] \frac{\overline{w}}{n} ,$$

are digitally representable. This will be the case, for example, if $\bar{a}$ is a power of 2 and $n = 2^s$ and if we take the integer p on pages 73-74 above, such that $p \leq N - s$. In order to guarantee this inequality we will assume that $N \geq s + p$. This done, we can define $\bar{y}_0^{(n)} = \bar{y}_0$ and for $i = 1, 2, \ldots, n$ with $\bar{x} \in S$ (digital numbers)

$$\bar{y}_i^{(n)}(\bar{x}) = \bar{A}(\bar{x}, 0, 1) \cap \left\{ \bar{y}_{i-1}^{(n)} \oplus \bar{Q}_i^{(n)}(\bar{x}) \right\},$$

$$\bar{y}_i^{(n)} = \bar{y}_i^{(n)}(x_i^{(n)}) ,$$

where $\bar{Q}_i^{(n)}(x)$ is defined by finite induction as follows:

$$\bar{Q}_0 = \bar{F}^{(k-1)} \left(x_i^{(n)}, \bar{A}(x_i^{(n)}, x_{i-1}^{(n)}, \bar{y}_{i-1}^{(n)}) \right) ,$$

and for $j = 0, 1, \ldots, k-1$,

$$\bar{Q}_{j+1} = \left\{ (\bar{Q}_j \odiv (k - j)) \otimes (\bar{x} \ominus x_{i-1}^{(n)}) \right\} \oplus \bar{F}^{(k-2-j)}(x_i^{(n)}, \bar{y}_i^{(n)}) ,$$

again with

$$\bar{F}^{(-1)}(x_i^{(n)}, \bar{y}_i^{(n)}) = \bar{y}_i^{(n)} .$$

Then

$$\bar{Q}_i^{(n)}(x) = \bar{Q}_k .$$

For rational $\overline{F}^{(\ell)}$, there exists an $M > 0$ such that for all $x \in \overline{X}^*$ (by (6.15) and the discussion of this section),

$$w(\overline{y}^{(n)}(\overline{x})) \leq w(y^{(n)}(\overline{x})) + w(\overline{y}^{(n)}(\overline{x}) - y^{(n)}(\overline{x}))$$

$$\leq M(\frac{\overline{w}}{n})^k + n\overline{M}\, 2^{-N}$$

(assuming $w(\overline{y}_o) \leq M_o(\frac{1}{n})^k + nM_o\, 2^{-N}$), where $\overline{y}^{(n)}(\overline{x}) = \overline{y}_i^{(n)}(\overline{x})$ for $\overline{x} \in X_i^{(n)} \bigcap S$. The quantities M and $\overline{M}$ depend on k and on the functions $\overline{F}^{(\ell)}$, $F^{(\ell)}$, $\ell = 0, 1, \ldots, k-1$. For fixed k, the integers n and N can be chosen large enough to make $w(\overline{y}^{(n)}(x))$ arbitrarily small for all $x \in \overline{X}^*$.

Choosing particular values for k, n, N, the digital interval version of the k^{th} order method will produce, by machine computation, intervals $\overline{y}_i^{(n)}$ containing values of the solution to $y' = f(x, y)$, $y(0) = y_o$ at certain values $x_i^{(n)}$, $i = 1, 2, \ldots, n$, of x. In fact,

$$y(x_i^{(n)}) \in \overline{y}_i^{(n)}, \qquad i = 1, 2, \ldots, n.$$

For intermediate values of x, we have

$$y(x) \in y^{(n)}(x) = y_i^{(n)}(x), \qquad x \in X_i^{(n)},$$

with $y_i^{(n)}(x)$ given by (6.14). Together with the relations

$$y_i^{(n)} \subset \overline{y}_i^{(n)}, \qquad i = 1, 2, \ldots, n,$$

this serves to bound $y(x)$ for all $x \in \overline{X}^*$, since we may replace $y_{i-1}^{(n)}$ by $\overline{y}_{i-1}^{(n)}$ in (6.14) and the resulting function on $X_i^{(n)} \left(- \{x_i^{(n)}\}\right)$ bounds $y_i^{(n)}(x)$. (With loss of continuity at the points $x_i^{(n)}$, however.)

A final question of importance in practice and difficult to answer in a simple way concerns an efficient choice of the integers n, k, N.

For the infinite precision interval methods, we found for the equation $y' = y^2$ with the initial condition $y(0) = 1$, that to maintain $w(y^{(n)}(x)) \leq 10^{-10}$ for $x \in [0, 1/4]$, the number of arithmetic operations was approximately minimized by choosing $k = 12$, $n = 5$. The number of operations was taken as nk^2 and it was assumed that

$$\sup_{x \in X^*} w(y^{(n)}(x)) = C(k+1)\left(\frac{1}{2n}\right)^k$$

for some constant C. Suppose now we make the assumption that the "contribution" to

$$\sup_{\overline{x} \in \overline{\overline{X}}^*} w(\overline{y}^{(n)}(\overline{x})) ,$$

made by replacing exact interval arithmetic by digital interval arithmetic is $nk^2\, 2^{-N}$. In fact, suppose that

$$\sup_{\overline{x} \in \overline{X}^*} w(\overline{y}^{(n)}(\overline{x})) = (k+1)\left(\frac{1}{2n}\right)^k + nk^2\, 2^{-N} .$$

Again, let us find integers k, n depending now on N such that nk^2 is minimum subject to the condition that

$$b(k, n, N) = (k+1)\left(\frac{1}{2n}\right)^k + nk^2\, 2^{-N} \leq 10^{-10} .$$

Clearly, we cannot satisfy the above inequality unless N is large enough; in particular, we must have

$$2^{-N} < 10^{-10} .$$

On the other hand, if N is sufficiently large we will again arrive at the result $k = 12$, $n = 5$. In fact, if N is such that

$$nk^2 \, 2^{-N} \leq \frac{1}{2} 10^{-10} ,$$

then the choice is not much affected, and for $k = 12$, $n = 5$ this means if $N > 43$, then $k = 12$, $n = 5$ is also close to the best choice for the digital version of the k^{th} order methods for this example.

The mechanization of some procedure enabling the computer to determine a good choice of $n, k,$ given a particular differential system, is certainly an interesting possibility.

The choice of the intervals B and a (see Figure 10) is arbitrary and actual machine computation (as well as some theoretical investigation) indicates the possibility of a reasonably efficient mechanizable procedure for choosing the successive "box" sizes, for a particular differential system.

Using single precision floating point digital interval arithmetic and the methods described above, actual machine programs were written for the IBM 7090 computer. Numerical results were obtained for a variety of differential systems including the equations for the restricted three-body problem. By going to high values of the order k surprisingly rapid computations produced very narrow rigorous bounds on the exact solution.

For instance, for the equation $y' = y$ with $y(0) = 1$ and $X^* = [0, 1]$, the 12^{th} order method $(k = 12)$ produced in one integration step $(n = 1)$ a machine interval containing e of width one binary digit in the last place carried, i.e., maximum single precision accuracy.

For the equation $y' = y^2$ with $y(0) = 1$, the 10^{th} order method with $n = 4$ and $X^* = [0, 1/4]$, produced an interval containing the exact solution $y(1/4) = 4/3$ of width less than one in the eighth significant decimal digit. Twenty-seven binary digits was full single precision accuracy; i.e., about .74 in the eighth significant decimal digit corresponds to one binary digit in the last place carried.

REFERENCES

[1] Burkill, J. C., "Functions of Intervals," Proc. London Math. Soc., 22 (1924), pp. 275-336.

[2] Burkill, J. C., "The Derivatives of Functions of Intervals," Fundamental Math. 5 (1924), pp. 321-327.

[3] Dwyer, P. S., Linear Computations, New York, Wiley (1951).

[4] Dwyer, P. S., "Errors of Matrix Computations," Simultaneous Equations and Eigenvalues, National Bureau of Standards, Applied Mathematics Series 29 (1953), pp. 49-58.

[5] Fischer, P. C., "Automatic Propagated and Round-Off Error Analysis," paper presented at the 13th national meeting of the Association for Computing Machinery, June 1958. Preprints available from ACM headquarters, 2 East 63rd Street, New York 21, N. Y.

[6] Gorn, Saul, and Moore, R., "Automatic Error Control - The Initial Value Problem in Ordinary Differential Equations," Aberdeen Proving Ground, Maryland, Ballistic Research Laboratories Report No. 893, March 1953.

[7] Gorn, Saul, "The Automatic Analysis and Control of Computing Errors," J. Soc. Indust. Appl. Math. 2, (1954) 69-81.

[8] Henrici, P., "Discrete Variable Methods in Ordinary Differential Equations," Wiley, New York (1962).

[9] Householder, A. S., "Principles of Numerical Analysis," New York, McGraw-Hill (1953).

[10] Lowans, A. N., Davids, N., and Levenson, A., "Table of the Zeros of the Legendre Polynomials of Order 1-16 and the Weight Coefficients for Gauss' Mechanical Quadrature Formula," Mathematical Tables and Other Aids to Computation (1943).

[11] Milne-Thompson, L. M., "The Calculus of Finite Differences," Mac Millan, London (1933).

[12] Moore, R. E., "Automatic Error Analysis in Digital Computation," LMSD-48421, Lockheed Missile and Space Company (1959).

[13] Moore, R. E. and Yang, C. T., "Interval Analysis I," LMSD-285875, Lockheed Missiles and Space Company, Palo Alto, Calif. (1959).

[14] Moore, R. E., Strother, W., and Yang, C. T., "Interval Integrals," LMSD-703073, Lockheed Missile and Space Company (1960).

[15] von Neumann, J. and Goldstine, H., "Numerical Inversion of Matrices of High Order," Bull. American Mathematical Society 53 (1947).

[16] Riesz, F. and Sz.-Nagy, B., "Functional Analysis," Ungar, New York (1955), pp. 19ff.

[17] Saks, S., "Theory of the Integral," Warsaw (1937), pp. 165-169.

[18] Strother, W., "Continuity for Multi-Valued Functions and Some Applications to Topology" (doctoral dissertation), Tulane University (1952).

[19] Strother, W., "Fixed Points, Fixed Sets, and M-Retracts," Duke Math. J., Vol. 22, No. 4 (1955), pp. 551-556.

[20] Sunaga, T., "Theory of an Interval Algebra and Its Application to Numerical Analysis," RAAG Memoirs II, Gaukutsu Bunken Fukeyu-kai, Tokyo (1958).

[21] Young, R. C., "The Algebra of Many-Valued Quantities,"*/ Math. Annalen 104 (1931), pp. 260-290.

*/ See also Young's thesis for Ph.D. at Cambridge, 1929.

STANFORD UNIVERSITY

TECHNICAL REPORTS DISTRIBUTION LIST

CONTRACT Nonr-225(37)

(NR 044-211)

Argonne National Laboratory
Technical Information Division
9700 South Cass Ave.
Argonne, Illinois — 1

Armed Services Technical
Information Agency
Arlington Hall Station
Arlington 12, Virginia — 10

Chief of Naval Research
Code 432
Office of Naval Research
Washington 25, D. C. — 2

Commanding Officer
Office of Naval Research
Branch Office
1000 Geary Street
San Francisco 9, California — 1

Commanding Officer
Office of Naval Research Branch Office
495 Summer Street
Boston 10, Massachusetts — 1

Commanding Officer
Office of Naval Research
Branch Office, 346 Broadway
New York 13, New York — 1

Commanding Officer
Office of Naval Research
Branch Office
Navy No. 100, Fleet Post Office
New York, New York — 2

Commanding Officer
Office of Naval Research
Branch Office
1030 East Green Street
Pasadena 1, California — 1

Commanding Officer
U. S. Naval Electronics Laboratory
San Diego 52, California
Attn: R. F. Arenz
F. A. Sabransky — 1

Computer Reviews
P. O. Box 1184
Chapel Hill, North Carolina
Attn: Roger P. Hansen
Editorial Ass't — 1

Department of Mathematics
Harvard University
Cambridge, Massachusetts — 1

Department of Mathematics
Massachusetts Institute of Technology
Cambridge, Massachusetts — 1

Space Technology Laboratories
One Space Park
Redondo Beach, California
Attn: Irene Matthews — 1

Institute for Mathematical Sciences
New York University
New York 3, New York — 1

Librarian, Communications Research
Division
Institute for Defense Analyses
Von Neumann Hall
Princeton, New Jersey — 1

Office of Technical Services
Department of Commerce
Washington 25, D. C. — 1

Document Center
Stanford Research Institute
333 Ravenswood Avenue
Menlo Park, California — 1

Technical Information Officer
Naval Research Laboratory
Washington 25, D. C. — 6

The Mitre Corporation Library
Bedford, Massachusetts
Attn: Supervisor
Library Services — 2

University of California
Attn: Library, Numerical Analysis Research
Los Angeles 24, California — 1

U. S. Naval Ordnance Laboratory
White Oak
Silver Spring, Maryland
Attn: Mathematics Department — 1

U. S. Naval Ordnance Laboratory
White Oak
Silver Spring, Maryland
Attn: Library — 1

U. S. Navy Management Office
Department of the Navy
Washington 25, D. C. — 1

U. S. Naval Weapons Laboratory
Dahlgren, Virginia — 1

Professor W. F. Atchison
Rich Electronic Computer Center
Georgia Institute of Technology
Atlanta, Georgia — 1

Dr. Erwin H. Bareiss
Applied Mathematics Division
Argonne National Laboratory
9700 South Cass Avenue
Argonne, Illinois — 1

Professor Robert C. F. Bartells
Computing Center
University of Michigan
Ann Arbor, Michigan — 1

Dr. Richard Bellman
The RAND Corporation
1700 Main Street
Santa Monica, California — 1

Professor Garrett Birkhoff
Mathematics Department
Harvard University
Cambridge 38, Mass. — 1

Bernard H. Bissinger
Department of Math., Math Library
Lebanon Valley College
Annville, Pennsylvania — 1

Mr. Truman Botts
Department of Mathematics
University of Virginia
Charlottesville, Virginia — 1

Professor A. T. Brauer
Department of Mathematics
University of North Carolina
Chapel Hill, North Carolina — 1

Professor George W. Brown
Western Data Processing Center
University of California
Los Angeles 24, California — 1

Professor John W. Carr, III
Mathematics Department
University of North Carolina
Chapel Hill, North Carolina — 1

Mr. Robert L. Causey
Lockheed Aircraft Corporation
Missiles and Space Division
3251 Hanover Street
Palo Alto, California — 1

Professor Abraham Charnes
The Technological Institute
Northwestern University
Evanston, Illinois — 1

Professor B. A. Chartres
Division of Applied Mathematics
Brown University
Providence 12, Rhode Island — 1

Professor Harvey Cohn
Department of Mathematics
University of Arizona
Tucson 25, Arizona — 1

Professor Stephen Crandall
Department of Mechanics
Mass. Inst. of Technology
Cambridge 39, Massachusetts — 1

Dr. Philip Davis
Division 11
National Bureau of Standards
Washington 25, D. C. — 1

Professor David B. Dekker
Director, Research Computer Lab.
University of Washington
Seattle 5, Washington — 1

Professor R. J. DeVogelaere
Department of Mathematics
University of California
Berkeley 4, California — 1

Professor C. L. Dolph
311 Park Avenue
Ann Arbor, Michigan — 1

Professor Jim Douglas, Jr.
Department of Mathematics
The Rice Institute
Houston 1, Texas — 1

Professor R. J. Duffin
Mathematics Department
Carnegie Institute of Technology
Pittsburgh 13, Pennsylvania — 1

Dr. Robert Dickson
Lockheed Aircraft Corporation
Missiles and Space Division
3251 Hanover Street
Palo Alto, California — 1

Professor Paul Dwyer
403 Lenawee Drive
Ann Arbor, Michigan — 1

Mr. Thomas A. Dwyer
Director, Computer Science
Department of Physics
University of Dayton
Dayton 9, Ohio — 1

Mr. Edgar L. Eichhorn
Burroughs Corporation
Electrodata Facility
460 Sierra Madre Villa
Pasadena, California — 1

Mr. Frank Engel, Jr.
Analytical Department 4-L-39
Westinghouse Electric Corp.
700 Braddock Avenue
East Pittsburgh, Pennsylvania — 1

Dr. G. W. Evans, II
Stanford Research Institute
Menlo Park, California — 1

Professor F. A. Ficken
Department of Mathematics
New York University
University Heights
New York 53, New York — 1

Mr. Werner Frank
4617 Loarkwood Avenue
Woodland Hills, California — 1

Professor J. N. Franklin
California Institute of Technology
Pasadena 4, California — 1

Professor Bernard Friedman
Department of Mathematics
University of California
Berkeley 4, California — 1

Dr. H. H. Germond
RCA Service Company
Bldg. 989-1 MU 811
Patrick Air Force Base
Florida 1

Professor Wallace Givens
Department of Engineering Science
Northwestern University
Evanston, Illinois 1

Dr. H. H. Goldstine
Research Center
International Business Machines
Yorktown, New York 1

Professor Saul Gorn
Moore School of Electrical Engineering
University of Pennsylvania
Philadelphia 4, Pennsylvania 1

Professor L. Greenberg
Department of Mathematics
Brown University
Providence, Rhode Island 1

Professor R. T. Gregory
Computation Center
University of Texas
Austin 12, Texas 1

Dr. T. N. E. Greville
5011 Strathmore Avenue
Kensington, Maryland 1

Professor Preston Hammer
Director, Numerical Analysis Lab.
University of Wisconsin
Madison 6, Wisconsin 1

Dr. Eldon Hansen
347 Aldean
Mt. View, California 1

Professor T. J. Higgins
Director, University of Wisconsin
Engineering Experiment Station
Madison 6, Wisconsin 1

Professor T. W. Hildebrandt
Department of Mathematics
Ohio State University
Columbus, Ohio 1

Dr. William Hooker
Philco Corporation
3875 Fabian Way
Palo Alto, California 1

Dr. A. S. Householder
Oak Ridge National Lab.
P.O. Box X
Oak Ridge, Tennessee 1

Dr. M. A. Hyman
IBM Systems Development Center
7220 Wisconsin Ave.
Bethesda, Maryland 1

Professor Eugene Isaacson
Institute of Mathematical Sciences
New York University
4 Washington Place
New York 3, New York 1

Professor Thomas A. Keenan
Computing Center
University of Rochester
Rochester 20, New York 1

Professor William B. Kehl
Computation and Data Processing Center
825 Cathedral of Learning
University of Pittsburgh
Pittsburgh 13, Pennsylvania 1

Dr. Ervand G. Kogbetliantz
Rockefeller Inst. for Medical Research
66th St. and York Ave.
New York, N. Y. 1

Professor R. Kodis
Department of Mathematics
Brown University
Providence, Rhode Island 1

Professor C. Donald La Budde
Department of Engineering Sciences
The Technological Institute
Northwestern University
Evanston, Illinois 1

Professor R. E. Langer
Department of Mathematics
University of Wisconsin
Madison 6, Wisconsin 1

Professor Arvid Lonseth
Department of Mathematics
Oregon State College
Corvallis, Oregon 1

Library
Digital Computer Lab.
University of Illinois
Urbana, Illinois 1

Mr. B. Mittman
Armour Research Foundation
10 West 35th Street
Chicago 16, Illinois 1

Professor James M. Moore
Dept. of Industrial Engineering
Northeastern University
Boston 12, Massachusetts 1

Professor T. S. Motzkin
Department of Mathematics
University of California
Los Angeles 24, California 1

Dr. J. P. Nash
Missile and Space Division
Lockheed Aircraft Corp.
Palo Alto, California 1

Professor C. D. Olds
P.O. Box 462
Los Altos, California 1

Dr. Joseph Otterman
P.O. Box 2008
University of Michigan
Ann Arbor, Michigan 1

Professor Alan J. Perlis
Computation Center
Carnegie Institute of Technology
Pittsburgh 13, Pennsylvania 1

Professor Clay L. Perry
Director, Computation Center
University of California
School of Science and Engineering
La Jolla, California 1

Professor Murray Protter
Department of Mathematics
University of California
Berkeley 4, California 1

Professor R. D. Richtmyer
AEC Computing and Applied Math. Center
Institute of Mathematical Sciences
New York University
4 Washington Place
New York 3, New York 1

Dr. Samuel Schechter
Research Scientist
AEC Computing Center
New York University
New York 3, New York 1

Professor Lowell Schoenfeld
Mathematics Department
Pennsylvania State University
State College, Pennsylvania 1

Professor Thomas H. Southard
Chairman, Mathematics Department
Alameda State College
22300 Foothill Blvd.
Hayward, California 1

Professor M. L. Stein
Department of Mathematics
Institute of Technology
University of Minnesota
Minneapolis 4, Minnesota 1

Professor J. J. Stoker
Institute of Mathematical Sciences
25 Waverly Place
New York 3, New York 1

Mr. Roland Thompson
Burroughs Corporation
6071 2nd Ave.
Detroit 32, Michigan 1

Dr. Gene Thompson
United Technology Corp.
1050 E. Arques Ave.
Sunnyvale, California 1

Professor John Todd
Mathematics Department
Calif. Inst. of Technology
Pasadena 4, California 1

Professor Richard S. Varga
Department of Mathematics
Case Institute of Technology
University Circle
Cleveland 6, Ohio 1

Professor Wolfgang Wasow
Department of Mathematics
University of Wisconsin
Madison 6, Wisconsin 1

Mr. H. C. Thacher, Jr.
Argonne National Laboratory
9700 South Cass Avenue
Argonne, Illinois 1

Professor David M. Young, Jr.
Department of Mathematics
University of Texas
Austin 12, Texas 1

Dr. James M. Ortega
Division 5426
Sandia Corporation
Albuquerque
New Mexico 1

<u>Distribution via ONR London</u>

Commanding Officer
Branch Office
Navy No. 100
Fleet Post Office
New York, New York

Mathematisch Centrum
2e Boerhaavestraat 49
Amsterdam 0, NETHERLANDS 1

Professor A. C. Aitken
Mathematical Institute
University of Edinburgh
16 Chambers Street
Edinburgh 1, SCOTLAND 1

Dr. Enzo Aparo
Instituto Nazionale per le
Applicazioni del Calcolo
Piazzale del Scienze 7
Roma, ITALY 1

Professor Dr. Charles Blanc
Centre de Calcul Electronique
Avenue de Cour, 33
Lausanne, SWITZERLAND 1

Professor Dr. F. L. Bauer
Institut für Angewandte Mathematik
Universität Mainz
Jakob-Welder-Weg 7
Mainz-Rhein
GERMANY 1

Mr. Niels Ivar Bech
Regnecentralen
Danish Institute for Computing
Machinery
Gl. Carlsbergvej 2
Valby, DENMARK 1

Dr. Stig Comét
International Computation Centre
Palazzo degli Uffici
Zona dell'E. U. R.
Roma, Italy 1

Contract Nonr-225(37)
October, 1962

Professor Dr. L. Collatz
Institut für Angewandte Mathematik
Universität Hamburg
(24a) Hamburg 13
Rothenbaumchaussee 67/69
GERMANY 1

Dr. Leslie Fox
Oxford University
Oxford, ENGLAND 1

Dr. Carl-Erik Fröberg
Avd. för Numerisk Analys
Lunds Universitet
Lund, SWEDEN 1

Mr. Richard Goodman
Director, Automatic Programming
Information Center
Department of Mathematics
The Technical College
Brighton, ENGLAND 1

Dr. J. M. Hammersley
Institute of Statistics
St. Cross Road
Oxford, ENGLAND 1

Professor Dr. Peter Henrici
Lehrstuhl für höhere Mathematik
Eigenössische technische hochschule
Zurich, SWITZERLAND 1

Dr. Joseph Hersch
Institut Battelle
7, route de Drize
Carouge, Geneva
SWITZERLAND 1

International Computation Center
Palazzo degli uffici
zona dell'E.U.R.
Roma, ITALY 1

Professor J. Kuntzmann
Laboratoire de Calcul
Universite de Grenoble
44-46 Avenue Felix-Viallet
Grenoble, Isére
FRANCE 1

Professor Pentti Laasonen
Finland Institute of Technology
Helsinki, FINLAND 1

Professor C. Lanczos
Institute for Advanced Studies
64-65 Merrion Square
Dublin, IRELAND 1

Professor L. Mirsky
The University
Sheffield, ENGLAND 1

Dr. A. R. Mitchell
Mathematics Department
St. Salvator's College
University of St. Andrews
St. Andrews, Fife, SCOTLAND 1

Mr. Jean A. Mussard
Executive Secretary
Provisional International
Computation Center
Palazzo degli Uffici
Zona dell'E.U.R.
Roma, ITALY 1

Professor A. Ostrowski
Certenago-Montagnola
bei Lugano
SWITZERLAND 1

Professor Dr. R. Sauer
Mathematisches Institut
der Technischen Hochschule
München 2
Arcisstr. 21, GERMANY 1

Professor E. Stiefel
Institut für angewandte Mathematik
Eidg. Technische Hochschule
Zurich, SWITZERLAND 1

Professor Dr. Heinz Unger
Institut für angewandte Mathematik
der Universität Bonn
Wegelerstrasse 10
Bonn, GERMANY 1

Professor Dr. A. Walther
Direktor, Institut für Praktische
Mathematik
Technische Hochschule
(16) Darmstadt, GERMANY 1

M. S. B. L. Wilson
Mathematics Department
Royal College of Science and
Technology
George Street
Glasgow C1, SCOTLAND 1

Professor Rigal
Secretaire a l'AFCAL
48 rue des Poitiers
Montbeliard, Doubs
FRANCE 1

Professor Dr. H. W. Wielandt
Mathematisches Institut
Universität Tübingen
Tübingen (14b)
Wilhelmstrasse 7, GERMANY 1

Dr. M. V. Wilkes
University Mathematical Lab.
Corn Exchange Street
Cambridge, ENGLAND 1

Dr. J. H. Wilkinson
Mathematics Division
National Physical Lab
Teddington, Middlesex
ENGLAND 1

Other Foreign Addresses

Dr. John M. Bennett
Adolph Basser Computing Lab.
University of Sydney
Sydney, N.S.W., AUSTRALIA 1

W. A. Cowan, Librarian
Barr Smith Library
The University
Adelaide
SOUTH AUSTRALIA 1

Professor A. Gonzalez Dominguez
The Institute of Calculus
University of Buenos Aires
Avenida De Mayo 760, 2º Piso
Buenos Aires, ARGENTINA 1

Facultad de Ciencias Exactas y Naturales
Biblioteca
Sección Canje
Casilla de Correo 1766
Buenos Aires
REPUBLICA ARGENTINA 1

Professor Santiago Friedmann
Director del Centro de Computacion
Universidad de Chile
Casilla 2777
Santiago, CHILE 1

Professor Allan Gibb
Department of Mathematics
University of Alberta
Calgary, Alberta
CANADA 1

Dr. C. C. Gotlieb
Computation Center
McLennan Laboratory
University of Toronto
Toronto, CANADA 1

Dr. J. F. Hart
Head, Computing Centre
The University of Western Ontario
London, CANADA 1

Professor Thomas E. Hull
Computing Centre
University of British Columbia
Vancouver 8, B.C.
CANADA 1

Professor W. Kahan
Department of Mathematics
University of Toronto
Toronto, CANADA 1

Professor J. McNamee
Department of Mathematics
University of Alberta
Alberta, CANADA 1

Professor Manuel Sadosky
Instituto de Calculo
Facultad de Ciencias Exactas
y Naturales
Peru 27, Buenos Aires
REPUBLICA ARGENTINA 1

Professor H. Schwerdtfeger
Department of Mathematics
McGill University
Montreal, Quebec
CANADA 1

Mr. J. C. Vella
Publishing Secretary
Sociedad Argentina de Calculo
Avenido de Mayo 760 - 2º Piso
Buenos Aires, ARGENTINA 2

Additional copies for future
requirements 50

Contract Nonr-225(37)
October, 1962