JOURNAL OF COMPLEXITY **14**, 302–318 (1998)
ARTICLE NO. CM980478

On the Complexity of Self-Validating Numerical Integration and Approximation of Functions with Singularities

Knut Petras

Institut für Angewandte Mathematik, Technische Universität Braunschweig, Pockelsstrasse 14, Braunschweig D-38106, Germany
E-mail: k.petras@tu-bs.de

Received June 30, 1997

DEDICATED TO THE MEMORY OF PROFESSOR GÜNTHER HÄMMERLIN

We consider the complexity of numerical integration and piecewise polynomial approximation of bounded functions from a subclass of $C^k([a,b]\setminus Z)$, where Z is bounded by $O(\varepsilon^{-1/k})$ and we may have much higher costs. The situation changes if we also allow "realistic" estimates of ranges of functions or derivatives on intervals as observations. A very simple algorithm now yields an error less than ε with $O(\varepsilon^{-1/k})$-costs and an analogous result is also obtained for uniform approximation with piecewise polynomials. In a practical implementation, estimation of ranges may be done efficiently with interval arithmetic and automatic differentiation. The cost for each such evaluation (also of ranges of derivatives) is bounded by a constant times the cost for a function evaluation. The mentioned techniques reduce the class of integrands, but still allow numerical integration of functions from a wide class with $O(\varepsilon^{-1/k})$ arithmetical operations and guaranteed precision ε.

1. MOTIVATION

One main point in the dispute between B. N. Parlett (1992) on the one hand, and J. F. Traub and H. Woźniakowski (1992) on the other, about the relevance of information-based complexity (IBC) was the problem class F. Let $I: F \to G$ be a mapping into the normed linear space G. IBC theory is interested in the estimation of the minimal cost, which is necessary to approximate $I[f]$ for $f \in F$ with at most an error ε. "The ingredient of IBCT that allows it to generate irrelevant results is the problem class F. ... We might say that knowledge of membership in F is information and

0885-064X/98 $25.00

should have a cost attached to it" (B. N. Parlett, 1992). "Do researchers in other disciplines charge for F? For example, researchers in numerical analysis often analyze the cost and error of important algorithms. The analysis depends on F. To give a simple example, the analysis of the composite trapezoidal rule usually requires that the second derivative of the integrand is bounded. There is no charge for membership in F. Indeed, how would one charge for knowing that a function has a bounded second derivative?" (J. F. Traub and H. Woźniakowski, 1992).

Only reliable information about the membership to the class F is able to produce a reliable output. Although we do not agree with B. N. Parlett's opinion that "whenever F is very large, it is realistic to assign no cost it" (How can the size be measured?), there are situations, where the membership is obvious and situations, where it may be "expensive" to verify. In general, it seems to be difficult to find a bound on a higher derivative of f but it seems to be relatively easy to recognize that a function f is composed from standard functions and that it may have only singularities of, e.g., the type $|x-z_0|^\alpha$, $\alpha>0$, and $z_0 \in \mathbb{R}$ fixed, or jump discontinuities produced by a piecewise definition of f. In this second case, we will see that it is not very expensive to extract information about the ranges of higher derivatives, which is useful for error estimation. However, what is also important, we may use this more global information in order to make the algorithm faster in a certain precise sense. The reason for that seems to be that adaption (respectively mathematics) can be more effective when it is based on additional "reliable global information" instead of only "local information" such as function evaluations.

2. INTRODUCTION

In many applications, the numerical calculation of integrals is desired or it is one constituent in a more complex problem. Consequently, there are many proposals for linear quadrature formulae Q_m of the form

$$Q_m[f] = \sum_{\nu=1}^{m} a_\nu f(x_\nu) \tag{1}$$

or

$$Q_m[f] = \sum_{\nu=1}^{m} a_\mu f^{(\nu_\mu)}(x_\mu),$$

or for adaptive quadrature schemes using at most m function evaluations or evaluations of derivatives (see, e.g., the books of Brass (1977), Davis

and Rabinowitz (1984), or Krommer and Überhuber (1994)). We define the remainder functional R_m by

$$R_m[f] = \int_{-1}^{1} f(x)\,dx - Q_m[f].$$

If we are sure that the integrand f is sufficiently smooth, we might trust the output of most of the frequently used quadrature methods in "many" cases, since the rate of convergence of either sequences of linear quadrature formulae with increasing number of nodes x_μ or of adaptive techniques is high. The situation changes if there are singularities, possibly hidden in any derivative, whose locations and types are unknown. It is not possible to guarantee the detection of singularities by using only function values or values of derivatives (see Wasilkowski and Gao (1992)). If a singularity is not detected, it may influence the speed of convergence drastically. While functions in $A^k[a, b]$, i.e., functions with a bounded kth derivative, may be integrated with an error of order $O(n^{-k})$ by using n function values, we can expect an error of the order $n^{-\alpha-1}$ for functions similar to $f(x) = |x - x_0|^\alpha$ (no linear method, yielding $o(n^{-\alpha-1})$ for all x_0, is known). This is a particular great decline if $\alpha \ll k$. However, the main result in Wasilkowski and Gao's article is that, in a certain sense, singularities can be found with high probability.

EXAMPLE. The standard routine for our purposes, DQAGS, of the standard integration package, QUADPACK (see Piessens *et al.* (1983)), is used to calculate

$$I_z := \int_0^1 \left(\sin x + \frac{1}{8} |x - z|^{3/2} \right) dx$$

for different values of z. First we require a relative accuracy of 10^{-8} and choose $z = z_i = (2i-1)/400$, $i = 1, ..., 100$. In most cases, DQAGS calculates I_z with the given accuracy, while in two cases, the error is more than 100 times the permitted error. Requiring an accuracy of 10^{-9} and choosing $z = z_i = (2i-1)/4000$, $i = 1, ..., 1000$, there is one z such that the error is more than 10,000 times the permitted error.

Of course, it depends on the underlying problem, whether errors as in the example are acceptable. We cannot avoid such errors if we may only use function values of the integrand or of some of its derivatives. If the values of the integrand themselves stem from the solution of a complicated numerical problem, we can indeed use only function values and we have to live with possible uncertainties in the output of an integration routine.

The situation changes, if we may allow "realistic range estimates" of f or its derivatives as observations. We first define the set of functions.

DEFINITION 1. We say that a function f is in $S_k[a, b]$ if it is bounded on the interval $[a, b]$ and if there is a subset $Z := Z(f) = \{z_0, ..., z_r\}$ of $[a, b]$ and a constant c such that f has a kth derivative on $[a, b]\backslash Z$ satisfying

$$|f^{(k)}(x)| \leqslant c \cdot \operatorname{dist}(\{x\}, \tilde{Z})^{-k}, \qquad \text{where} \quad \tilde{Z} := Z \cup \{a-1\}.$$

Here and in the following, for arbitrary subsets Y_1, Y_2 of $\mathbb{R}^d$, $d \geqslant 1$, we use the notation

$$\operatorname{dist}(Y_1, Y_2) = \inf\{\|y_1 - y_2\|_\infty \mid y_i \in Y_i\}.$$

(The set $\tilde{Z}$ is defined for the case $Z = \varnothing$).

What could "realistic range estimates" for functions in $S_k[a, b]$ mean? In practice, it may be difficult to estimate f on the whole interval $[a, b]$ in one step (see Section 4), but the corresponding mapping should at least yield a bound for the range of f if we ask for it on a sufficiently small subinterval of $[a, b]$. Furthermore, if $\operatorname{dist}(\{z_i\}, [c, d])$ is much smaller than $d-c$, then it may be too difficult to estimate $f^{(k)}$ on $[c, d]$. We therefore require that a "centered extension" of $[c, d]$ by a certain factor $\beta \geqslant 1$ still does not contain z_i. By a centered extension of an interval $[x-y, x+y]$, we mean

$$\beta \diamond [x-y, x+y] := [x-\beta y, x+\beta y].$$

For the next definition we need the modulus of an interval $[c, d]$,

$$|[c, d]| := \max\{|c|, |d|\}.$$

DEFINITION 2. We say that the pair (E_0, E_k) is a realistic range estimator for $f \in S_k[a, b]$ if the following conditions are satisfied.

(i) There are positive numbers δ and M such that $E_0(f, [c, d]) \supset f([c, d])$ is an interval, whose width is less than M whenever $[c, d] \subset [a, b]$ and $d-c < \delta$;

(ii) There are numbers $\beta \geqslant 1$ and γ such that for each $[c, d] \subset [a, b]$, $E_k(f, [c, d])$ is an interval (possibly of infinite width), which contains the set $f^{(k)}([c, d])$ and which satisfies

$$|E_k(f, [c, d])| \leqslant \gamma \operatorname{dist}(\beta \diamond [c, d], Z)^{-k}.$$

The intervals $E_i(f, [c, d])$, $i = 0, k$, are then called realistic range estimates.

In Section 3, we describe a simple algorithm for integration of $f \in S_k[a, b]$ involving function evaluations and realistic range estimates. For all such functions, we can prove that we obtain the precision ε after $O(\varepsilon^{-1/k})$ function evaluations and range estimations. Modifying the set of functions, we show an analogous result for uniform approximation in Section 4.

If the integrand is given as a combination of standard or library functions, we have all necessary information to produce realistic range estimates and therefore to integrate numerically with an error within any required accuracy. Although it does not seem to be well known ("global information of that kind is only rarely available in practice" [anonymous]) interval arithmetic and differentiation tools such as automatic differentiation are able to provide realistic range estimators in many cases very efficiently. Basic properties of these two techniques are therefore described in Section 5. In Section 6, we discuss the consequences of the practical computation of realistic range estimates for the application of the results stated in Sections 3 and 4.

3. AN ALGORITHM FOR SELF-VALIDATING NUMERICAL INTEGRATION

In order to formulate our algorithm, we first have to introduce some notation and notions from quadrature theory.

Of course, we may take any quadrature formula Q_m of the form (1) and transform it to any interval $[c, d]$ (see Brass (1977, p. 14)). We denote this transformed formula by $Q_{m;[c,d]}$. Let $R_m[\mathbb{P}_k] = 0$, where $\mathbb{P}_k$ denotes the set of all polynomials of degree less than or equal to k. Denote furthermore by $\operatorname{Var} g$ the total variation of the function g. Then, the Peano kernel representation (see Brass (1977, p. 41))

$$R_m[f] = \int_{-1}^{1} f^{(s)}(x)\, dK_{s+1}(x)$$

$$= \int_{-1}^{1} \left[f^{(s)}(x) - \frac{1}{2}\left(\max_t f^{(s)}(t) + \min_t f^{(s)}(t)\right) \right] dK_{s+1}(x),$$

$$s = 0, 1, \ldots, k$$

yields constants

$$c_s = 2^{-s-2} \operatorname{Var} K_{s+1},$$

such that

$$|R_{m;[c,d]}[f]| \leqslant c_s \cdot (c-d)^{s+1} \cdot \max_{x,\, y \in [c,d]} |f^{(s)}(x) - f^{(s)}(y)|,$$

where $s = 0, 1, ..., k$.

Suppose now that $f \in S_k[a, b]$ and that we have a realistic range estimator (E_0, E_k). We obtain

$$|R_{m;[c,d]}[f]| \leqslant \min_{i \in \{0,k\}} c_i \cdot (c-d)^{i+1} \cdot E_i(f, [c, d]) =: u(Q_m, f, [c, d]),$$

i.e., we may say that $u(Q_m, f, [c, d])$ is the uncertainty of Q_m for the function f on the interval $[c, d]$, (with respect to (E_0, E_k)).

ALGORITHM 1. Our initial (sub-)interval is $[a, b]$.

(1) Estimate the uncertainties on each new subinterval.

(2) If the sum of the uncertainties is less than the required accuracy, apply the quadrature formula Q_m on each subinterval and stop. Otherwise subdivide the interval with the greatest uncertainty in two intervals of equal width and continue with step (1).

THEOREM 1. *Let $f \in S_k[a, b]$ and let (E_0, E_k) be a realistic range estimator. Then, Algorithm 1 yields an ε-approximation to $\int_a^b f(x)\, dx$ after at most $O(\varepsilon^{-1/k})$ range estimations and function evaluations.*

The proof of the theorem requires the following estimate.

LEMMA. *Let $D_0 > 0$, $\varrho > 1$ and let $\alpha > 0$. Define*

$$D_{j+1} = D_j \left\{ 1 + \alpha \left(\frac{D_0}{D_j} \right)^{1/\varrho} \right\}.$$

Then, there is a constant

$$q \geqslant \min \left\{ 1, \left(\frac{2\alpha}{\varrho 2^{\varrho}} \right)^{\varrho} \right\}$$

such that

$$D_j > q \cdot j^{\varrho} \cdot D_0.$$

Proof of the Lemma. We prove the lemma by induction over j. The lemma obviously holds for $j=1$ if $q<1+\alpha$, since

$$D_1=(1+\alpha)\,D_0>q\cdot D_0.$$

By the induction hypothesis, we obtain

$$D_{j+1}>q\cdot j^{\varrho}\cdot D_0\cdot\left(1+\frac{\alpha}{j\cdot q^{1/\varrho}}\right).$$

This is greater than or equal to $q\cdot(j+1)^{\varrho}\cdot D_0$ if and only if

$$\eta_j:=\frac{(1+1/j)^{\varrho}-1}{1/j}<\frac{\alpha}{q^{1/\varrho}}.$$

Defining $h(x)=(1+x)^{\varrho}$, we see that

$$\eta_j\leqslant\max_{x\in[0,\,1]}|h'(x)|=h'(1)=\frac{\varrho}{2}\cdot 2^{\varrho},$$

i.e., the number q only has to satisfy

$$\varrho 2^{\varrho-1}\leqslant\frac{\alpha}{q^{1/\varrho}}.$$

This yields the statement. ∎

Proof of the Theorem. First note that the properties of E_0 imply that the maximal uncertainty on the subintervals would tend to zero if the algorithm would have no exit point. For a given $\lambda>0$, we may therefore assume that, at a certain time, the algorithm has generated subintervals of $[a,b]$ such that the maximal uncertainty on any subinterval does not exceed λ for the first time. We will estimate how many subintervals may have been generated at most. Therefore, we have to estimate the widths of the produced subintervals from below. Let, almost without restriction,

$$[a_j,a_{j+1}]=[a_j,a_j+\sigma]\subset\left[z_\nu,\frac{z_\nu+z_{\nu+1}}{2}\right]$$

be one of the subintervals and let, for simplicity, $z_\nu=0$. Our interval originates from a bisection of $[a_j-\sigma,a_j+\sigma]$ or of $[a_j,a_j+2\sigma]$ such that the uncertainty on one of these two intervals must be greater than λ. This

means at least that the estimate using the kth derivative is greater than λ, i.e.,

$$\underbrace{c_k \gamma \cdot 2^{k+1}}_{=: L} \cdot \sigma^{k+1} \cdot (a_j - \beta\sigma)^{-k} > \lambda. \tag{2}$$

Setting

$$a_{j+1} - a_j = k_j \cdot a_j,$$

inequality (2) is equivalent with

$$a_{j+1} > a_j \left\{ 1 + \left(\frac{\lambda}{L a_j} (1 - \beta k_j)^k \right)^{1/(k+1)} \right\}.$$

In the case $k_j > 1/(\beta + 1)$, we trivially have

$$a_{j+1} > \left(1 + \frac{1}{\beta + 1} \right) a_j. \tag{3}$$

Otherwise, we obtain

$$a_{j+1} > a_j \left\{ 1 + \left(\frac{\lambda}{L(1+\beta)^k\, a_j} \right)^{1/(k+1)} \right\}. \tag{4}$$

The right-hand side in (4) is less than that in (3) if and only if

$$a_j > \frac{(1+\beta)\,\lambda}{L}. \tag{5}$$

Therefore, inequality (4) always holds under the assumption (5). Let now $[a_0, a_1], [a_1, a_2], \ldots$ be the consecutive subintervals to the right of $(1+\beta)\lambda/L$. Since the right-hand side of (4) is an increasing function of a_j, we obtain

$$a_j \geqslant D_j,$$

where

$$D_0 := \frac{(1+\beta)\,\lambda}{L}$$

and

$$D_{j+1} := D_j \left\{ 1 + \frac{1}{1+\beta} \left(\frac{D_0}{D_j} \right)^{1/(k+1)} \right\} \geqslant q \cdot (j+1)^{k+1} \cdot D_0.$$

The last inequality is an application of the lemma above, where we set $\alpha = 1/(1+\beta)$ and $\varrho = k+1$. We therefore have

$$a_j \geqslant \frac{z_\nu + z_{\nu+1}}{2} \qquad \text{if} \quad j > \left(\frac{z_\nu + z_{\nu+1}}{2qD_0} \right)^{1/(k+1)} = \text{const} \cdot \lambda^{-1/(k+1)}.$$

Suppose now that our algorithm has produced the subinterval $[c, c+\sigma]$. This was done by a bisection of an interval of width 2σ. The uncertainty on this latter interval must have been greater than λ. By assumption (i) of the theorem,

$$\sigma \geqslant \frac{\lambda}{2M} \qquad \text{or} \qquad \sigma \geqslant \frac{\delta}{2},$$

where the constants M and δ are introduced in the theorem. Therefore, the number of produced subintervals that intersect $[0, (1+\beta)\,\lambda/L]$ is at most $1 + (1+\beta)\,\lambda/(L \cdot \min\{\delta/2, \lambda/(2M)\})$, which is bounded from above by a fixed constant if $\lambda \leqslant 1$.

In summary, we see that $[z_\nu, (z_\nu + z_{\nu+1})/2]$ and thus also the whole interval $[a, b]$ is completely covered by $O(\lambda^{-1/(k+1)})$ produced subintervals. We need $O(\lambda^{-1/(1+k)})$ calculations of function values and of ranges of the function and its derivatives to obtain the accuracy $\lambda \cdot \lambda^{-1/(1+k)}$. This is exactly what had to be proved, since each subdivision costs two calculations of ranges. ∎

4. AN ALGORITHM FOR SELF-VALIDATING NUMERICAL APPROXIMATION

The problem of approximating a given function uniformly on an interval $[a, b]$ has a more local character than numerical integration. This makes the formulation of approximation algorithms principally easier. We stop refining our approximation in a certain subinterval exactly if we have obtained there the given precision. The results, however, require more stringent assumptions. Boundedness of the given function by a constant M is not sufficient, since it can only yield boundedness of the approximation error by M. We need a smoothness assumption. Here, we choose

Lipschitz continuity, i.e., boundedness of the first derivative, but this is only exemplary.

DEFINITION 1'. We say that a function f is in $\tilde{S}_k[a, b]$ if its first derivative is bounded on the interval $[a, b]$ and if there is a subset $Z := Z(f) = \{z_0, ..., z_r\}$ of $[a, b]$ and a constant c such that f has a kth derivative on $[a, b] \backslash Z$ satisfying

$$|f^{(k)}(x)| \leqslant c \cdot \operatorname{dist}(\{x\}, \tilde{Z})^{-k-1}, \qquad \text{where} \quad \tilde{Z} := Z \cup \{a-1\}.$$

Realistic range estimators for the class $\tilde{S}_k[a, b]$ clearly have to satisfy assumptions different from those for $S_k[a, b]$.

DEFINITION 2'. We say that the pair (E_0, E_k) is a realistic range estimator for $f \in \tilde{S}_k[a, b]$ if the following conditions are satisfied.

(i) There are positive numbers δ and M such that $E_0(f, [c, d]) \supset f([c, d])$ is an interval, whose width is less than $M(d-c)$ whenever $[c, d] \subset [a, b]$ and $d-c < \delta$.

(ii) There are numbers $\beta \geqslant 1$ and γ such that for each $[c, d] \subset [a, b]$, $E_k(f, [c, d])$ is an interval (possibly of infinite width), which contains the set $f^{(k)}([c, d])$ and which satisfies

$$|E_k(f, [c, d])| \leqslant \gamma \operatorname{dist}(\beta \diamondsuit [c, d], Z)^{-k+1}.$$

The intervals $E_i(f, [c, d])$, $i = 0, k$, are again called realistic range estimates.

In order to formulate our approximation algorithm, we first have to choose a standard projection $L: C[-1, 1] \to \mathbb{P}_k$, for example, an interpolation operator using $k+1$ nodes. For the error

$$R[f] = f - L[f]$$

we have again $R[\mathbb{P}_k] = 0$. Transformed on the interval $[c, d]$, the corresponding error $R_{[c, d]}$ satisfies error estimates

$$\|R_{[c, d]}[f]\| \leqslant c_s \cdot (d-c)^s \cdot \max_{x, y \in [c, d]} |f^{(s)}(x) - f^{(s)}(y)|,$$

$$\text{where} \quad s = 0, 1, ..., k. \tag{6}$$

Analogously to the integration problem, we define the uncertainty of the approximation operator on the interval $[c, d]$.

ALGORITHM 2. Our initial (sub-)interval is $[a, b]$.

(1) Estimate the uncertainties on each new subinterval.

(2) If the uncertainty on a subinterval is not less than the required accuracy, subdivide this interval in two intervals of equal width and continue with step (1).

(3) Transform and "apply" L on each of the subintervals.

Remarks. (1) The application of L in step (3) of the algorithm can mean, for example, that we calculate the coefficients of the approximation in terms of a basis of polynomials (e.g., in terms of transformed Chebyshev polynomials).

(2) One often requires some smoothness conditions for the approximation. If we use an interpolation operator with nodes at the boundary of each subinterval, we obtain continuity. We may, for example, use interpolation at the extrema of the Chebyshev polynomials of the first kind; this yields relatively small error constants c_s. If we want to have a higher degree of smoothness, we might use multiple nodes at the boundary. However, then Eq. (6) is not always satisfied for small s.

THEOREM 2. *Let $f \in \tilde{S}_k[a, b]$ and let (E_0, E_k) be a realistic range estimator for f in the sense of Definition* 2'. *Then, Algorithm* 2 *yields a piecewise polynomial approximation with uniform error less than ε after at most $O(\varepsilon^{-1/k})$ range estimations and function evaluations.*

The proof is almost the same as that of Theorem 1. Essentially the only difference is that we have to replace λ by ε and to replace each occurrence of k in an exponent by $k-1$. Then we obtain the accuracy ε by a subdivision into $O(\varepsilon^{-1/k})$ intervals.

5. TOOLS FOR SELF-VALIDATING NUMERICAL COMPUTATIONS

As indicated in the Introduction, interval arithmetic and automatic differentiation may form efficient realistic range estimators. For convenience, we list some basic facts about these techniques. Detailed descriptions may be found in the books of Alefeld and Herzberger (1983) and Rall (1981).

Due to the fact that computers can only try to approximate the mathematical reality and due to the necessity to guarantee that computed results

are reliable, interval arithmetic was developed. It enables an automatic error control. Interval arithmetic uses statements such as

$$\text{if} \quad x \in [a, b] \qquad \text{and} \qquad y \in [c, d], \quad \text{then} \quad x + y \in [a + c, b + d]$$
$$(\text{i.e., } [a, b] + [c, d] = [a + c, b + d])$$

or

$$\text{if} \quad x \in [a, b] \quad \text{then} \quad e^x \in [e^a, e^b] \quad (\text{i.e., } e^{[a, b]} = [e^a, e^b])$$

to guarantee that the output of a calculation is in a certain interval (possibly of infinite width) if the input values are in certain intervals. In a realization on a computer, we have to round down the left end-point of the resulting intervals and round up the right end-point in order to get guaranteed inclusions.

For the formulation of the assumptions below, we have to introduce the width

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

of an interval. One should be aware of the fact that elementary rules of calculus are not valid in each case. For example, denoting by sqr the square function, we readily verify

$$\text{sqr}([-1, 2]) = [0, 4] \qquad \text{but} \qquad [-1, 2] \cdot [-1, 2] = [-2, 4].$$

Different ways of representing a function may yield different outputs of an interval computation. This implies that we have to do mathematics not for functions but for their representations. In order to control an expression representing f, we replace the argument at each of its, say r, occurrences by a separate variable, for example,

$$f(x) = \frac{x}{x+1} \cdot \cos x \qquad \text{then} \quad e_f(x, y, z) = \frac{x}{y+1} \cdot \cos z$$

and

$$f(x) = e_f(x, x, x).$$

We see that there is no difference between using f or e_f when dealing only with numbers as arguments. For intervals X, interval arithmetic calculates $e_f(X^r)$ instead of $f(X)$ and of course, we have $f(X) \subset e_f(X^r)$. We may also

consider $f(X)$ as the range of e_f on the main diagonal of X^r. This illustrates how interval arithmetic may increase the range. However, if e_f is Lipschitz continuous, then the increase is bounded by a constant times $w(X)$ (see Alefeld and Herzberger, 1981, Chap. 3).

In particular removable singularities, e.g., in a ratio, usually cause trouble in interval computations since they cause output intervals of infinite width. If f is continuous this does not have to be true for its expression e_f. Using interval arithmetic we can therefore treat only subclasses of continuous functions, namely those with an expression $e_f(x, y, z, ...)$ being bounded on a certain region, which has the diagonal $x = y = z = ...$ in its interior. On the other hand, suppose that there is one simple zero in the numerator and one in the denominator of an expression. In order to apply any numerical integration or approximation routine reasonably, we should at least be able to decide whether both zeros coincide or not. This is not possible in general. If the zeros coincide, the singularity is removable and the numerical algorithm may yield a reasonable result. If they do not, the integral does not exist but the algorithm may yield a number, which is of course misleading. It is therefore not obvious that our restriction is a true one for a general purpose integration algorithm.

Finally we note that $e_f([a, b]^r) \cup e_f([b, c]^r) \subset e_f([a, c]^r)$ for $a \leqslant b \leqslant c$ and often, the difference between the two sides is considerable. A useful way to estimate the range of f more accurately on an interval is therefore successive subdivision of this interval. If e_f is bounded by a fixed constant, whenever its argument X^r at least satisfies $w(X) < \delta$ with a fixed positive δ, then we can obtain a global bound on f after a finite number of subdivisions.

The errors of many numerical algorithms for calculating a functional on $C^k[a, b]$ may be estimated in terms of ranges of derivatives. If we have a reasonable expression for derivatives, interval arithmetic may therefore yield explicit error bounds. In order to obtain such an expression, we could first start a computer algebra system and then differentiate the given expression a certain number of times. This would take a lot of time and the resulting expressions are often lengthy. A very effective alternative is the use of automatic differentiation (AD). AD simply replaces each operation with function values by an operation with vectors, which contain the numerical values of Taylor coefficients, according to certain rules (see Rall, 1981). Therefore we might use the same output of the compiler but a different arithmetic afterwards, showing that we have no additional cost for compilation. The cost for a calculation of the kth derivative is bounded by a fixed constant $d_k = O(k^2)$, which is independent of the given function, times the cost for the evaluation of a function value.

Of course, AD implicitly also uses an expression, say $e_{f^{(s)}}: \mathbb{R}^{r_s} \to \mathbb{R}$, for the derivative $f^{(s)}$ but usually a relatively economic one. Furthermore, it

may easily be combined with interval arithmetic. This combination is thus an effective way to calculate ranges of derivatives on intervals. Now, a measure for the complexity is therefore the number of function calls, no matter whether a function value, or a vector of estimates for Taylor coefficients is desired.

6. THE PRACTICAL COMPUTATION OF REALISTIC RANGE ESTIMATES AND ITS CONSEQUENCES

As indicated in the preceding section, for a given interval $[c, d]$, we may apply interval arithmetic to get an estimate for $f([c, d])$ and we may combine interval arithmetic and AD in order to estimate $f^{(k)}([c, d])$. From Theorem 1, we readily obtain

COROLLARY 1. *Let* $f \in S_k[a, b]$, *denote by* $e_f: [a, b]^r \to \mathbb{R}$ *the used expression for* f, *and by* $e_{f^{(k)}}([a, b]\backslash Z)^{r_k} \to \mathbb{R}$ *the used expression for* $f^{(k)}$. *Furthermore, assume that*

(i) e_f *is bounded in a neighbourhood of the main diagonal of* $[a, b]^r$,

(ii) *there are constants* $\beta \geqslant 1$ *and* γ *such that*

$$|e_{f^{(k)}}(X^{r_k})| \leqslant \gamma \operatorname{dist}(\beta \diamondsuit X, Z)^{-k},$$

for all intervals X *satisfying* $(B \diamondsuit X) \cap Z = \varnothing$.

Then, the pair (E_0, E_k) *defined by*

$$E_0(f, X) = e_f(X^r) \qquad \text{and} \qquad E_k(f, X) = e_{f^{(k)}}(X^{r_k}) \tag{7}$$

is a realistic range estimator and Algorithm 1 *yields an* ε*-approximation of* $\int_a^b f(x)\,dx$ *after at most* $O(\varepsilon^{-1/k})$ *arithmetic operations.*

Remarks. (1) The assumption (ii) may be interpreted in such a way that there is a region $G \subset \mathbb{R}^{r_k}$ (see Fig. 1 for $r_k = 2$) on which $e_{f^{(k)}}(x)$, $x = (\xi_1, ..., \xi_{r_k}) \in \mathbb{R}^{r_k}$ is defined and bounded by $\text{const} \cdot \operatorname{dist}(\{x\}, \partial G)^{-k}$, where ∂G is the boundary of G. The region G is the intersection of nonempty circular cones with vertices at the points $(z_i, ..., z_i)$ and with the main diagonal $\xi_1 = \cdots = \xi_{r_k}$ as symmetry axis.

(2) If we want to estimate $f^{(k)}$ via interval arithmetic and AD, we obtain estimates for lower order derivatives automatically or with almost

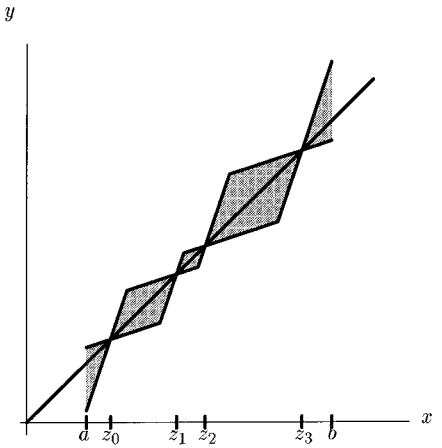

FIGURE 1

no additional costs. It may therefore improve the algorithm by operating with a modified uncertainty

$$\tilde{u}(Q_m, f, [c, d]) = \min_{0 \leqslant s \leqslant k} c_s (d-c)^{s+1} w(e_{f^{(s)}}([c, d]^{r_s})),$$

where $r_0 = r$ and r_s is defined analogously to r_k.

(3) The algorithm is of course similar to many other proposals of self-validating integration algorithms (see, e.g., Corliss, 1987, or Corliss and Rall, 1987).

(4) Surprisingly, for the functions considered in Theorem 1, the asymptotically most costly part of the algorithm is the search for the interval with maximal uncertainty. For example, suppose we have a list of μ intervals ordered by uncertainty and we have to insert a produced subinterval. This takes $\sim \ln \mu$ comparisons. Producing n subintervals takes $\sim n$ function evaluations but $\sim \sum_{\mu=1}^{n} \ln \mu \sim n \ln n$ comparisons. We should therefore modify the algorithm such that we collect the intervals with uncertainty $\in [2^k, 2^{k+1})$, $k \in \mathbb{Z}$ in a class J_k and that we do not always have to subdivide the interval with maximal uncertainty, but any interval from J_k with maximal k. Since the uncertainty on an interval I is at least twice the uncertainty on an interval produced by bisection of I, it can be seen that the number of function evaluations of the modified algorithm is at most twice that of the original algorithm.

Of course, we also obtain

COROLLARY 2. *Let* $f \in \tilde{S}_k[a, b]$, *denote by* $e_f: [a, b]^r \to \mathbb{R}$ *the used expression for* f *and by* $e_{f^{(k)}}: ([a, b] \backslash Z)^{r_k} \to \mathbb{R}$ *the used expression for* $f^{(k)}$. *Furthermore, assume that*

(i) e_f *is Lipschitz continuous in a neighbourhood of the main diagonal of* $[a, b]^r$,

(ii) *there are constants* $\beta \geqslant 1$ *and* γ *such that*

$$|e_{f^{(k)}}(X^{r_k})| \leqslant \gamma \operatorname{dist}(\beta \diamondsuit X, Z)^{-k-1},$$

for all intervals X *satisfying* $(\beta \diamondsuit X) \cap Z = \varnothing$.

Then, the pair (E_0, E_k) *defined by* (7) *is a realistic range estimator and Algorithm* 2 *yields a piecewise polynomial approximation with uniform error less than* ε *after at most* $O(\varepsilon^{-1/k})$ *arithmetic operations.*

Analogous remarks as above, except for Remark 3, also concern Corollary 2.

ACKNOWLEDGMENTS

The author thanks the referees for helpful comments and Erich Novak for an extensive and fruitful discussion both concerning arrangement and presentation of the paper.

REFERENCES

Alefeld, G., and Herzberger, J. (1983), "Introduction to Interval Computations," Academic Press, New York.

Brass, H. (1977), "Quadraturverfahren," Vandenhoeck & Ruprecht, Göttingen.

Corliss, G. F. (1987), Computing narrow inclusions for definite integrals, *in* "Computer-arithmetric" (E. Kaucher, U. Kulisch, and C. Ullrich, Eds.), pp. 150–169, Teubner, Stuttgart.

Corliss, G. F., and Rall, L. B. (1987), Adaptive, self-validating numerical quadrature, *SIAM J. Sci. Statist. Comput.* **8**, 831–847.

Davis, P. J., and Rabinowitz, P. (1984), "Methods of Numerical Integration," Academic Press, London.

Krommer, A. R., and Überhuber, C. W. (1994), "Numerical Integration on Advanced Computer Systems," Lecture Notes in Comput. Sci., Vol. 848, Springer-Verlag, Berlin.

Parlett, B. N. (1992), Some basic information on information-based complexity theory, *Bull. Amer. Math. Soc.* (*N.S.*) **26**, 3–27.

Pressens, R., de Doncker-Kapenga, E., Überhuber, C. W., and Kahaner, D. K. (1983), "QUADPACK," Springer Ser. Comput. Math., Vol. 1, Springer-Verlag, Berlin.

Rall, L. B. (1981), "Automatic Differentiation—Techniques and Applications," Lecture Notes in Comput. Sci., Vol. 120, Springer-Verlag, Berlin.

Traub, J. F., and Woźniakowski, H. (1992), Perspectives on information-based complexity, *Bull. Amer. Math. Soc.* (*N.S.*) **26**, 29–52.

Wasilkowski, G. W., and Gao, F. (1992), On the power of adaptive information for functions with singularities, *Math. Comp.* **58**, 285–304.