JOURNAL OF APPROXIMATION THEORY 47, 195–202 (1986)

# Quadrature and Widths

ERICH NOVAK

*Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstraße 1 1/2, D-8520 Erlangen, West Germany*

*Communicated by Oved Shisha*

Received March 2, 1984

## SUMMARY

The $n$-widths of a subset $A$ of a Banach space $B$ describe how well the elements of $A$ can be approximated by elements of $n$-dimensional subspaces of $B$. This paper investigates relations between $n$-widths and error estimates for quadrature formulas. These estimates describe how well a linear form on a function class $A$ can be approximated by quadrature formulas with $n$ knots. For the case that $A$ is a class of bounded functions, we compare the $n$-widths $d_n(A)$, referring to the sup-norm, with deterministic error estimates $e_n(A, \mu)$, for some linear form $\mu$ on $A$, defined by

$$e_n(A, \mu) = \inf_{\mu_n} \sup_{f \in A} |\mu(f) - \mu_n(f)|,$$

where $\mu_n$ runs through all quadrature formulas with $n$ knots. In order to get from $e_n(A, \mu)$ to a quantity which depends only on $A$, we introduce

$$e_n(A) = \sup_{\|\mu\| \leqslant 1} e_n(A, \mu).$$

For an arbitrary class $A$ of bounded functions, the following relation to the $n$-widths holds:

$$e_n(A) \leqslant 2 \cdot d_n(A).$$

In spite of this connection, the behavior of $d_n$ and $e_n$ turns out to be very different, in general. For instance, the asymptotic behavior of $d_n$ and $e_n$ is the same if $A$ is a Hölder class or a Sobolev class with $p \geqslant 2$ but is different if $A$ is a Sobolev class with $1 \leqslant p < 2$. The last statement will be shown in another paper, by a different method.

In the second part of the paper, we investigate stochastic error bounds $\sigma_n(A, \mu)$ and $\sigma_n(A)$ for stochastic quadrature formulas, introduced via variances, and their relations to the $n$-widths, based on the $L_2$-norm. We



0021-9045/86 \$3.00

improve a general lower estimate of Bahvalov for $\sigma_n(A, \mu)$ and give new stochastic error bounds for some special function classes. Concerning the asymptotic behavior, we see that in some interesting cases the stochastic error bounds converge faster than the deterministic ones. Quantitatively, the improvement amounts to the factor $n^{-1/2}$.

## 1. Deterministic Quadrature Formulas

Let $X$ be an arbitrary set, $B(X) = \{f: X \to \mathbb{R} \mid f \text{ bounded}\}$, and $A \subseteq B(X)$. First we give a result which generalizes the well-known interpolation theorem (see Shapiro [14]):

Proposition 1. *Let $V \subseteq B(X)$ be a vector space with* $\dim V = n \in \mathbb{N}$ *and $L: V \to \mathbb{R}$ linear. Then for each $\varepsilon > 0$ there exist $x_1, \ldots, x_n \in X$ and $a_1, \ldots, a_n \in \mathbb{R}$ with*

$$L(f) = \sum_{i=1}^{n} a_i f(x_i) \text{ for all } f \in V$$

*and*

$$\|L\| \leqslant \sum_{i=1}^{n} |a_i| \leqslant \|L\| + \varepsilon.$$

*Proof.* In the case where $X$ is compact and $V$ consists of continuous functions, this proposition holds even for $\varepsilon = 0$ (this is the interpolation theorem mentioned). We reduce the general case to this special case: The set $M = \{f \in V \mid \|f\| = 1\}$ is compact, and for $\delta_1 > 0$ there exist $f_1, \ldots, f_m \in M$ with $M = \bigcup_{i=1}^{m} \{f \in M \mid \|f - f_i\| \leqslant \delta_1\}$. For $\delta_2 > 0$ let $x_i$ be given so that $|f_i(x_i)| \geqslant 1 - \delta_2$ $(i = 1, \ldots, m)$. By eventually making $K' = \{x_1, \ldots, x_m\}$ larger we can assume that for $V' = \{f/K \mid f \in V\}$ with $K$ finite and $K' \subseteq K$ the statement $\dim V' = n$ holds.

Now we apply the above-mentioned special case to $K$, $V'$, and $L'$ defined by $L'(f/K) = L(f)$. Because of $\|f/K\| \geqslant \|f\| \cdot (1 - \delta_1 - \delta_2)$ for all $f \in V$ we get $L(f) = \sum_{i=1}^{n} a_i f(x_i)$ with

$$\|L\| \leqslant \sum_{i=1}^{n} |a_i| = \|L'\| \leqslant \|L\| \cdot (1 - \delta_1 - \delta_2)^{-1}$$

and the statement follows.

Now we define the $n$th error bounds for deterministic quadrature formulas that we want to compare with the $n$-widths, referring to the sup-norm:

DEFINITION. Let $\mu \in B'(X)$ (the latter being the dual of $B(X)$) and $M_n = \{\mu_n \in B'(X) \mid \mu_n(f) = \sum_{i=1}^n a_i f(x_i)$ for some $a_i$ and $x_i\}$. Set

$$e_n(A, \mu) = \inf_{\mu_n \in M_n} \sup_{f \in A} |\mu(f) - \mu_n(f)|$$

and

$$e_n(A) = \sup_{\|\mu\| \leq 1} e_n(A, \mu).$$

The $n$-widths of $A$ in $B(X)$ as defined by Kolmogorov [8] are given by

$$d_n(A) = \inf_{X_n} \sup_{f \in A} \inf_{g \in X_n} \|f - g\|,$$

where $X_n$ runs through all vector spaces of $B(X)$ of dimension $n$.

PROPOSITION 2.

$$e_n(A, \mu) \leq 2 \cdot \|\mu\| \cdot d_n(A)$$

*and therefore*

$$e_n(A) \leq 2 \cdot d_n(A).$$

*Proof.* Let $A$, $\mu$, $n$, and $\varepsilon > 0$ be given. There is a vector space $V \subseteq B(X)$ with $\dim V = n$ and $\sup_{f \in A} \inf_{g \in V} \|f - g\| \leq d_n(A) + \varepsilon$. Because of Proposition 1 there is a $\mu_n \in M_n$ with $\mu_n(f) = \mu(f)$ for all $f \in V$ and $\|\mu_n\| \leq \|\mu\| + \varepsilon$. Therefore $|\mu_n(f) - \mu(f)| \leq (d_n(A) + \varepsilon) \cdot (\|\mu\| + \|\mu_n\|) \leq (d_n(A) + \varepsilon) \cdot (2\,\|\mu\| + \varepsilon)$ for all $f \in A$. The statement follows from this.

*Remark.* By examples one can show that the constant 2 in Proposition 2 is optimal; i.e., in general it cannot be replaced by a smaller constant.

EXAMPLES. (1) Let $X$ be compact, $A \subseteq C(X)$ and $M(X) = C'(X)$ the set of all (Radon-) measures on $X$.

Then

$$e_n(A) = \sup_{\substack{\mu \in M(X) \\ \|\mu\| \leq 1}} e_n(A, \mu).$$

(2) Let $C^{k,\alpha}([0,1]^s)$ ($s \in \mathbb{N}$, $k \in N_0$, $0 < \alpha \leq 1$) be the Hölder class $\{f\colon [0,1]^s \to \mathbb{R} \mid |D^{(k)}f(x) - D^{(k)}f(y)| \leq \|x - y\|^\alpha$ for all derivatives of order $k\}$. Then $e_n(C^{k,\alpha}([0,1]^s), \lambda^s) \asymp e_n(C^{k,\alpha}([0,1]^s) \asymp d_n(C^{k,\alpha}([0,1]^s)) \asymp n^{-(k+\alpha)/s}$ holds. Here $a_n \asymp b_n$ means that there exist $c_1, c_2 > 0$

with $c_1 < a_n/b_n < c_2$ for all $n \in \mathbb{N}$. This result is a consequence of Proposition 2 and known facts about $e_n(C^{k,\alpha}([0, 1]^s), \lambda^s)$ (see Bahvalov [1]) and $d_n(C^{k,\alpha}([0, 1]^s))$ (see Lorentz [10] or Tihomirov [15]).

PROBLEM. We state the following problem: For which $A \subseteq B(X)$ does $e_n(A) \to 0$? The analogous problem for $d_n(A)$ is easy: $d_n(A) \to 0$ iff $A \subseteq V + S$ with a finite-dimensional $V$ and a compact $S$. As Proposition 3c shows, there are "very large" sets $A \subseteq B(X)$ with $e_n(A) \to 0$.

We now investigate $e_n(A)$ for a vector space $A$. The analogous problem for $d_n(A)$ is very easy: $d_n(A) = \infty$ holds for all $n < \dim A$ and $d_n(A) = 0$ for all $n \geq \dim A$. We consider the case where $X$ is a compact space and $A \subseteq C(X)$:

PROPOSITION 3. (a) *Let $X$ be compact and scattered (the latter means that there is no nonempty subset of $X$ without isolated points; see Semadeni [13]) and $A \subseteq C(X)$,* $\dim A = m$. *Then $e_n(A) = \infty$ holds for all $n < m$.*

(b) *Let $X$ be compact and $A \subseteq C(X)$,* $\dim A > n$, *and let $A$ contain a Chebyshev system of $n$ functions. Then $e_n(A) = \infty$ holds.*

(c) *If $X$ is compact but not scattered then there is an $A \subseteq C(X)$ with* $\dim A = \infty$ *but $e_n(A) = 0$ for all $n \in \mathbb{N}$. That means that for each $\mu \in B'(X)$ there is a representation*

$$\mu(f) = a_0 f(x_0) \qquad (\textit{for all } f \in A).$$

*Proof.* (a) Let $A = \langle f_1, ..., f_m \rangle$ and $\dim A = m \geq 2$ (there is nothing to prove for $m = 1$). Then $Y = \{(f_1(x), ..., f_m(x)) \mid x \in X\} \subseteq \mathbb{R}^m$ is countable (see Semadeni [13]). Because $\mathbb{R}^m$ is not the union of countable many subspaces of dimension $n < m$, the set $M = \{y \in \mathbb{R}^m \mid y = \sum_{i=1}^n \lambda_i y_i,\ \lambda_i \in \mathbb{R},\ y_i \in Y\}$ is a proper subset of $\mathbb{R}^m$. Let $\tilde{y} \in \mathbb{R}^m \setminus M$ and consider a $\mu \in B'(X)$ with $\tilde{y} = (\mu(f_1), ..., \mu(f_m))$. It is easy to show that $e_n(A, \mu) = \infty$ is valid and therefore the statement follows.

(b) Let $A = \langle f_1, ..., f_m \rangle$ with $\dim A = m > n$ and let $\{f_1, ..., f_n\}$ be a Chebyshev system. Let $\mu \in B'(X)$ with $\mu(f_i) = 0$ for $i = 1, ..., n$ and $\mu(f_{n+1}) = 1$. Assuming $\mu(f) = \sum_{i=1}^n a_i f(x_i)$ for all $f \in A$ (where we can presume the $x_i$ to be different) from $\sum_{i=1}^n a_i f_j(x_i) = 0$ for $j = 0, ..., n$, it follows that $a_i = 0$ for all $i = 1, ..., n$, which contradicts the fact that $\sum_{i=1}^n a_i f_{n+1}(x_i) = 1$.

(c) Because $X$ is compact but not scattered, there exists a continuous $h: X \to [0, 1]$ which is onto. Because $[-1, 1]^{\mathbb{N}}$ is a Peano space there even exists a continuous $h^*: X \to [-1, 1]^{\mathbb{N}}$ which is onto. Then the projections $h_i^*(i \in \mathbb{N})$ of $h^*$ are continuous and linear independent. Therefore, for $A = \langle f_1, f_2, ... \rangle$, $\dim A = \infty$ holds. Let $\mu \in B'(X)$ with $\|\mu\| \leq 1$. Then

$\mu(h_i^*) = k_i \in [-1, 1]$ for all $i \in \mathbb{N}$. There is an $x_0 \in X$ with $h^*(x_0) = (k_1, k_2, \ldots)$. Then $\mu(h_i^*) = h_i^*(x_0)$ holds for all $i$ and therefore $e_1(A, \mu) = 0$ and the statement follows.

*Remark.* For the Sobolev classes $W_p^k([0,1]^s) = \{f\colon [0,1]^s \to \mathbb{R} \mid \sum_{|\alpha|=k} \|D^{(\alpha)}f\|_p \leq 1\}$ in the case $pk > s$ (imbedding condition)

$$e_n(W_p^k([0,1]^s), \lambda^s) \asymp e_n(W_p^k([0,1]^s)) \asymp n^{-k/s}$$

holds (see Novak [11]).

The lower estimate follows from arguments similar to Bahvalov [1], and the upper estimate in the case $2 \leq p \leq \infty$ follows from Proposition 2 and known facts about the $n$-widths of these classes. The case $1 \leq p < 2$ is much more difficult because then the relation

$$e_n(W_p^k([0,1]^s)) \asymp d_n(W_p^k([0,1]^s))$$

is not valid (see Höllig [6] or Kashin [7]).

## 2. Stochastic Quadrature Formulas

Let $(X, \mathfrak{a}, \mu)$ be a finite signed measure space and $A$ a set of $\mu$-integrable functions on $X$. A stochastic quadrature formula $Q_n \in S_n$ with $n$ knots is a random variable with values in $X^n \times \mathbb{R}^n = M_n(X)$. By $Q_n(f)$ we denote the random variable

$$Q_n(f) = \sum_{i=1}^{n} a_i f(x_i), \qquad \text{where } Q_n = (x_1, \ldots, x_n, a_1, \ldots, a_n).$$

Analogously to the $e_n(A, \mu)$ we now define the $n$th error bound for stochastic quadrature formulas,

$$\sigma_n(A, \mu) = \inf_{Q_n \in S_n} \sup_{f \in A} (E(\mu(f) - Q_n(f))^2)^{1/2}$$

(where $E$ is the expectation of a random variable). For a given measurable space $(X, \mathfrak{a})$ we define

$$\sigma_n(A) = \sup \sigma_n(A, \mu),$$

where $\mu$ runs through all signed measures on $(X, \mathfrak{a})$ with $\|\mu\| \leq 1$.

Example. If $X$ is compact and $\mathfrak{a}$ is the Baire $\sigma$-algebra then for $A \subseteq C(X)$ the numbers $e_n(A)$ and $\sigma_n(A)$ are directly comparable because the signed measures on $(X, \mathfrak{a})$ and the Radon measures on $C(X)$ correspond to each other.

The following statement is a more precise version of a result of Bahvalov [1]:

PROPOSITION 4. *Let* $A \subseteq L_1(X, \mathfrak{a}, \mu)$ *and* $f_i$ $(i = 1, \ldots, 2n)$ *with the following conditions*:

(i) *the* $f_i$ *have disjunct supports and fulfill* $\mu(f_i) \geqslant \varepsilon$ *for all* $i = 1, \ldots, 2n$,

(ii) *for all* $\delta_i \in \{-1, 1\}$ *the function* $\sum_{i=1}^{2m} \delta_i f_i$ *is an element of* $A$.

*Then* $\sigma_n(A, \mu) \geqslant (\varepsilon/2) \cdot n^{1/2}$ *is valid.*

*Remark.* Under the same conditions Bahvalov's [1] method gives $\sigma_n(A, \mu) \geqslant \varepsilon \cdot c \cdot n^{1/2}$ with some unfixed $c > 0$, independent of $n$.

*Proof.* Let $\tilde{A} = \{\sum_{i=1}^{2n} \delta_i f_i \mid \delta_i \in \{-1, 1\}\}$ and for $\mu_n \in M_n(X)$ let $F(\mu_n) = \sum_{f \in \tilde{A}} |\mu_n(f) - \mu(f)|^2$. Then

$$F(\mu_n) \geqslant 2^n \cdot \sum_{i=0}^{n} \binom{n}{i} \left(i - \frac{n}{2}\right)^2 \cdot \varepsilon^2 = 2^{2n} \cdot \varepsilon^2 \cdot \frac{n}{4}$$

is valid. Therefore, for all $Q_n \in S_n$ the relation $E(F(Q_n)) \geqslant 2^{2n} \cdot \varepsilon^2 \cdot (n/4)$ is valid and there exists an $f \in \tilde{A}$ with $E((Q_n(f) - \mu(f))^2) \geqslant \varepsilon^2 \cdot (n/4)$. From this the statement follows.

Now we compare the numbers $\sigma_n(A, \mu)$ with the $n$-width of $A$ in the space $L_2(X, \mathfrak{a}, \mu)$, which we denote with $d_{n,2}(A, \mu)$:

PROPOSITION 5. *Let* $\mu$ *be positive and* $A \subseteq L_2(X, \mathfrak{a}, \mu)$. Then $\sigma_{n+1}(A, \mu) \leqslant d_{n,2}(A, \mu) \cdot \|\mu\|^{1/2}$.

For the proof of Proposition 5 we need the following lemma, which is due to Ermakov and Zolotukhin [3]; see also Ermakov [2].

LEMMA. *Let* $\mu$ *be positive and* $A \subseteq L_2(X, \mathfrak{a}, \mu)$ *a vector space with* $\dim A = n$ *and* $1 \in A$. *Then there is a* $Q_n \in S_n$ *with the following properties*:

(a) $E(Q_n(f)) = \mu(f)$ *for all* $f \in L_1(X, \mathfrak{a}, \mu)$,

(b) $Q_n(f) = \mu(f)$ *for all* $f \in A$,

(c) $E((Q_n(f) - \mu(f))^2) \leqslant \|\mu\| \cdot \inf_{g \in A} \mu((f-g)^2)$ *for all* $f \in L_2(X, \mathfrak{a}, \mu)$.

*Proof of proposition* 5. For $\varepsilon > 0$ there is a linear space $V \subseteq L_2(X, \mathfrak{a}, \mu)$ with $\dim V = n$ and $\sup_{f \in A} \inf_{g \in V} \|f - g\|_2 \leqslant d_{n,2}(A, \mu) + \varepsilon$. We apply the lemma to the vector space $\langle V, 1 \rangle$ and get a $Q_{n+1} \in S_{n+1}$ with $\sup_{f \in A} E((Q_{n+1}(f) - \mu(f))^2) \leqslant \sup_{f \in A} \|\mu\| \cdot \inf_{g \in V} \mu((f-g)^2) \leqslant \|\mu\| \cdot (d_{n,2}(A, \mu) + \varepsilon)^2$ and the statement follows.

*Remark.* With the help of Proposition 5 and known estimates of $d_{n,2}(A, \mu)$ (see, for example, Korneicuk [9] and Parfenov [12]) one gets estimates for the $\sigma_n(A, \mu)$.

Now we want to compare the numbers $e_n(A, \mu)$ and $e_n(A)$ with the numbers $\sigma_n(A, \mu)$ and $\sigma_n(A)$, respectively:

PROPOSITION 6. (a) *Let $\mu$ be a signed finite measure on $(X, \mathfrak{a})$ and $A \subseteq L_1(X, \mathfrak{a}, \mu) \cap B(X)$. Then $\sigma_n(A, \mu) \leqslant e_n(A, \mu)$ holds.*

(b) *Let $X$ be compact, $\mathfrak{a}$ the Baire $\sigma$-algebra on $X$, and $A \subseteq C(X)$. Then $\sigma_n(A) \leqslant e_n(A)$ holds.*

*Remark.* This proposition is a simple consequence of the fact that each deterministic quadrature formula $\mu_n \in M_n$ can be regarded as a stochastic quadrature formula $Q_n \in S_n$ with constant $x_i$ and $a_i$. Stochastic quadrature formulas are interesting in those cases where the $\sigma_n(A, \mu)$ converge much faster than the corresponding $e_n(A, \mu)$.

Now we give some results for special function classes:

PROPOSITION 7. (a) *For the class $V = \{f\colon [0, 1] \to \mathbb{R} \mid \operatorname{Var}(f) \leqslant 1\}$ the statement $e_n(V) \asymp e_n(V, \lambda) \asymp \sigma_n(V) \asymp \sigma_n(V, \lambda) \asymp 1/n$ holds.*

(b) $\sigma_n(C^{k,\alpha}([0, 1]^s), \lambda^s) \asymp \sigma_n(C^{k,\alpha}([0, 1]^s)) \asymp n^{-(k+\alpha)/s - 1/2}$.

(c) $\sigma_n(W_p^k([0, 1]^s), \lambda^s) \asymp n^{-k/s - 1/2}$ *for* $p \geqslant 2$.

(d) $\sigma_n(W_p^k([0, 1]^s)) \asymp n^{-k/s - 1/2}$ *for* $kp > s$ *and* $p \geqslant 2$.

*Remarks.* (i) These results are due to the author [11] and contain those of several authors (see Bahvalov [1] and Haber [4, 5]).

(ii) The statement 7(a) shows that stochastic quadrature formulas do not always converge faster than deterministic ones. Another example would be the class $W_p^k([0, 1]^s)$ for $p = 1$.

(ii) In some cases the $\sigma_n(A, \mu)$ converge faster than the corresponding $e_n(A, \mu)$. This has been remarked for the Hölder classes in the case $\mu = \lambda^s$ by Bahvalov [1].

## ACKNOWLEDGMENTS

The results of this paper are part of the author's doctoral thesis [11]. The author would like to thank his supervisor, Professor D. Kölzow, for his initiation and valuable suggestions.

## REFERENCES

1. N. S. BAHVALOV, On approximate computation of integrals, *Vestnik MGV Ser. Math. Mech. Astronom. Phys. Chem.* **4** (1959), 3–18. [In Russian]

2. S. M. ERKAMOV, "Die Monte-Carlo-Methode und verwandte Fragen," Munich, 1975.
3. S. M. ERMAKOV AND V. G. ZOLOTUKHIN, Polynomial approximations and the Monte Carlo method, *Theory Probab. Appl.* **5** (1960), 428–431.
4. S. HABER, Numerical evaluation of multiple integrals, *SIAM Rev.* **12** (1970), 481–526.
5. S. HABER, Stochastic quadrature formulas, *Math. Comp.* **23** (1969), 751–764.
6. K. HÖLLIG, Diameters of classes of smooth functions, *in* "Quantitative Approximation: Proceedings, Symposium, 1979" (R. A. DeVore and K. Scherer, Eds.), pp. 163–175, Academic Press, New York, 1980.
7. B. S. Kashin, Diameters of some finite-dimensional sets and classes of smooth functions, *Math. USSR-Izv.* **11**, No. 2 (1977), 317–333.
8. A. N. KOLMOGOROV, Annäherung von Funktionen einer gegebenen Funktionenklasse, *Ann. of Math.* **31** (1936), 107–111.
9. N. P. KORNEICUK, Widths in $L_p$ of classes of continuous and of differentiable functions, and optimal methods of coding and recovering functions and their derivatives, *Math. USSR-Izv.* **18** (1982), 227–247.
10. G. G. LORENTZ, "Approximation of Functions," Holt, Rinehart & Winston, New York, 1966.
11. E. NOVAK, "Zur unteren Fehlergrenze von Quadraturverfahren," Doctoral thesis, Erlangen, 1983.
12. O. G. PARFENOV, Asymptotic behavior of the diameters of certain classes of analytic functions, *Functional Anal. Appl.* **15** (1982), 304–305.
13. Z. SEMADENI, "Banach spaces of continuous functions," Vol. 1, Warsaw, 1971.
14. H. S. SHAPIRO, "Topics in Approximation Theory," Lecture Notes in Mathematics, Vol. 187, Springer-Verlag, Berlin, 1971.
15. V. M. TIHOMIROV, Widths of sets in functional spaces and the theory of best approximations, *Russian Math. Surveys* **15** (1960), 75–112.