

IS GAUSS QUADRATURE OPTIMAL FOR  
ANALYTIC FUNCTIONS?

M.A. Kowalski, A.G. Werschulz and  
H. Wozniakowski

CUCS-81-83

IS GAUSS QUADRATURE OPTIMAL FOR  
ANALYTIC FUNCTIONS?

M.A. Kowalski  
Institute of Informatics  
University of Warsaw

A.G. Werschulz\*  
Division of Science and Mathematics  
Fordham University  
and  
Department of Computer Science  
Columbia University

H. Woźniakowski\*\*  
Institute of Informatics      and      Department of Computer Science  
University of Warsaw                      Columbia University

November, 1983

\* This research was supported in part by the National Science Foundation under Grants MCS-8203271 and MCS-8303111.

\*\* This research was supported in part by the National Science Foundation under Grant MCS-8923676.

Abstract

We consider the problem of optimal quadratures for integrands  $f: [-1,1] \rightarrow \mathbb{R}$  which have an analytic extension  $\bar{f}$  to an open disk  $D_r$  of radius  $r$  about the origin such that  $|\bar{f}| \leq 1$  on  $\bar{D}_r$ . If  $r = 1$ , we show that the penalty for sampling the integrand at zeros of the Legendre polynomial of degree  $n$  rather than at optimal points, tends to infinity with  $n$ . In particular there is an "infinite" penalty for using Gauss quadrature. On the other hand, if  $r > 1$ , Gauss quadrature is almost optimal. These results hold for both the worst-case and asymptotic settings.

## 1. Introduction.

This paper deals with approximations to  $\int_{-1}^1 f(x) dx$  by algorithms whose sole knowledge of  $f$  consists of samples at points from the interval  $[-1,1]$ . We assume that integrands belong to the class  $F(D_r)$  of functions  $f: [-1,1] \rightarrow \mathbb{R}$  having an analytic extension to  $D_r = \{z \in \mathbb{C}: |z| < r\}$  whose modulus is bounded by unity on  $\bar{D}_r$ .

One of the best-known methods of approximating such integrals is Gauss quadrature. This algorithm is derived by requiring the approximation to be exact for polynomials of as high a degree as possible. There are many papers dealing with the error analysis of Gauss quadratures, see for instance [1], [6], [7], [8], [9], [12]. In particular, sharp error estimates are known for analytic functions on the ellipse with foci  $\pm 1$  and the sum of semi-axes  $q$ , where  $q > 1$ . The behavior of Gauss quadrature for the class  $F(D_r)$  with  $r > 1$  follows easily. We know no previous results for the class  $F(D_1)$ .

The goal of this paper is not the study of Gauss quadrature per se. We are actually interested in the intrinsic error of Gauss information, i.e., the minimal error among all algorithms which evaluate the integrand at Gauss nodes.

(There is no a priori reason to believe that Gauss quadrature

uses Gauss information optimally, i.e., that the error of Gauss quadrature equals the intrinsic error of Gauss information.) Our aim is to compare the intrinsic error of Gauss information using  $n$  nodes to the  $n$ -th minimal error, i.e., the minimal error among all algorithms which evaluate the integrand at  $n$  points. When the former is worse than the latter, this tells us that not only is Gauss quadrature bad, but it is bad precisely because any algorithm using Gauss information is bad.

A number of papers [3],[10], [11] show that for large values of  $r$  Gauss quadrature is almost optimal. Note that as  $r$  increases the class  $F(D_r)$  looks more like a class of polynomials. while  $F(D_\infty)$  consists of constants. In this paper we address the question:

Is Gauss quadrature close to optimal for all  $r$ ?

We pursue our results in both a worst-case and an asymptotic setting [14], [15]. When  $r = 1$ , there is no "breathing room" between the interval of integration and the region of analyticity of integrands. We first consider the case  $r = 1$  in the worst-case setting. Due to [4], [5], the  $n$ -th minimal worst-case error is roughly  $\exp(-c\sqrt{n})$  for some  $c > 0$ . On the other hand, we show that in the worst-case setting the error of any algorithm using  $n$  Gauss nodes is at least about  $n^{-2}$ .

We next consider the case  $r = 1$  in the asymptotic setting. We apply general results of [15] to show that the  $n$ -th minimal asymptotic error is roughly  $\exp(-c\sqrt{n})$  for some  $c > 0$ . We prove that the minimal (asymptotic) error of any algorithm using  $n$  Gauss nodes is roughly  $n^{-2}$ .

Hence in both settings there is an unbounded penalty for using Gauss nodes when  $r = 1$ . We stress that this is a bad property of Gauss nodes rather than Gauss quadrature. That is, this holds for any algorithm using Gauss nodes.

We finally discuss the case  $r > 1$  which allows some "breathing room" between the interval of integration and the region of analyticity of integrands. We show that in both the worst-case and asymptotic settings the  $n$ -th minimal error is roughly  $\exp(-c n)$  for some  $c > 0$  and that Gauss quadrature is almost optimal.

Hence optimality of Gauss quadrature for analytic functions requires a strong assumption on analyticity of integrands  $f$ , i.e.,  $f \in F(D_r)$  with  $r > 1$ . The integration problem for the class  $F(D_r)$  with  $r > 1$  is essentially easier than the corresponding problem for the class  $F(D_1)$ .

## 2. How bad is Gauss quadrature when $r = 1$ ?

### 2.1. Worst-case setting.

We want to approximate  $\int_{-1}^1 f(x) dx$  for  $f \in F = F(D_1)$  using the following information about  $f$

$$N_n(f) = [f(x_1), f(x_2), \dots, f(x_n)]$$

where  $x_i \in [-1, 1]$ . By an algorithm  $\varphi$  we mean any mapping such that  $\varphi: N_n(F) \rightarrow \mathbb{R}$ . If  $\varphi$  is linear, i.e.,  $\varphi(N_n(f)) = \sum_{k=1}^n a_k f(x_k)$  for some  $a_k$ , we will refer to  $\varphi$  as a quadrature rule (or, more briefly, a quadrature).

In the worst-case setting we measure the error  $e(N_n, \varphi)$  of an algorithm  $\varphi$  using  $N_n$  by

$$e(N_n, \varphi) := \sup\left\{ \left| \int_{-1}^1 f(x) dx - \varphi(N_n(f)) \right| : f \in F \right\}.$$

It is well-known, see [2] and also [14, Thm. 3.1, p. 54], that

$$(1) \quad e(N_n) := \inf_{\varphi} e(N_n, \varphi) = \sup\left\{ \int_{-1}^1 f(x) dx : f \in F, N_n(f) = 0 \right\}$$

and the infimum in (1) is attained for a linear algorithm (quadrature). Bojanov [4], [5] proved that

$$(2) \quad e(n) := \inf_{x_1, x_2, \dots, x_n} e(N_n) \geq \exp(-5\pi\sqrt{n/2})$$

and found information  $N_n^*$  and a quadrature  $Q_n^*$ .

$$(3) \quad N_n^*(f) = [f(x_1^*), f(x_2^*), \dots, f(x_n^*)],$$

$$Q_n^*(N_n^*(f)) = \sum_{k=1}^n a_k^* f(x_k^*),$$

such that

$$(4) \quad e(N_n^*, Q_n^*) = \exp(-\pi\sqrt{n}/2).$$

Due to [2], see also [14, Thm. 7.1, p. 48], the estimate (2) is valid even if the points  $x_i$  are chosen adaptively, i.e.,  $x_i$  is allowed to depend on  $f(x_1), f(x_2), \dots, f(x_{i-1})$ . Hence  $N_n^*$  and  $Q_n^*$  are almost optimal, i.e., the choice of  $x_k^*$  and  $a_k^*$  nearly minimizes the error of any algorithm using  $n$  samples.

We compare the almost optimal information  $N_n^*$  with Gauss information  $N_n^G$ .

$$(5) \quad N_n^G(f) = [f(\zeta_1), f(\zeta_2), \dots, f(\zeta_n)],$$

where  $\zeta_k$  is the  $k$ -th zero of the Legendre polynomial  $P_n$ . We now prove that the choice of nodes  $\zeta_k$  is very poor.

Theorem 1:

$$e(N_n^G) = \Theta(n^{-2}).$$

Here we use the  $\Theta$ -notation, which may be thought of as a "two-sided"  $O$ -notation. That is,  $f = \Theta(g)$  iff  $f = O(g)$  and  $g = O(f)$ .

Proof: Let  $h(z) := \prod_{i=1}^n \frac{z - \zeta_i}{1 - z\zeta_i}$  ( $= \frac{P_n(z)}{z^n P_n(1/z)}$ ). Note that

$h \in F$ ,  $N_n^G(h) = 0$  and  $|h(z)| = 1$  for  $|z| = 1$ . We first prove that



$$(6) \quad \dots -I_n := \int_{-1}^1 |h(x)|^2 dx \leq e(N_n^G) \leq \int_{-1}^1 |h(x)| dx =: J_n.$$

Indeed, let  $N(f) := [f(\zeta_1), f'(\zeta_1), \dots, f(\zeta_n), f'(\zeta_n)]$ . Obviously  $e(N_n^G) \geq e(N)$ . Due to [4] we have  $e(N) = \int_{-1}^1 |h(x)|^2 dx$  which proves the left inequality of (6). Choose now an arbitrary  $g \in F$  such that  $N_n^G(g) = 0$ . Since the function  $g(z)/h(z)$  is analytic in  $D_1$ , by the maximum modulus principle we get

$$\sup_{z \in D_1} |g(z)/h(z)| = \sup_{|z|=1} |g(z)/h(z)| = \sup_{|z|=1} |g(z)| \leq 1.$$

Hence  $g(x) \leq |g(x)| \leq |h(x)|$  for every  $x \in [-1, 1]$ . Since  $g$  is arbitrary, this and (1) yield  $e(N_n^G) \geq J_n$ . The proof of (6) is completed.

By

$$P_n(\xi) = \pi^{-1} \int_0^\pi (\xi + \sqrt{\xi^2 + 1} \cos \theta)^n d\theta$$

(see [12, p. 87]), we get

$$\begin{aligned} x^n P_n(1/x) &= \pi^{-1} \int_0^{\pi/2} [(1 + t \cos \theta)^n + (1 - t \cos \theta)^n] d\theta \\ &= \pi^{-1} \sum_{k=0}^n \binom{n}{k} [t^k + (-t)^k] \int_0^{\pi/2} \cos^k \theta d\theta \end{aligned}$$

where  $t = \sqrt{1 - x^2}$  and  $x \in [0, 1]$ . For  $n \geq 4$  we have

$$\begin{aligned} (1 + \frac{(n-2)(n-3)}{8} t^2)^2 &\leq 1 + \frac{n(n-1)}{4} t^2 + \frac{n(n-1)(n-2)(n-3)}{64} t^4 \\ &\leq \pi^{-1} \sum_{k=0}^n \binom{n}{k} [t^k + (-t)^k] \int_0^{\pi/2} \cos^k \theta d\theta \leq \sum_{k=0}^n \binom{n}{k} t^k \\ &= (1 + t)^n \leq \exp(nt). \end{aligned}$$

We thus proved

$$(7) \quad \left[1 + \frac{(n-2)(n-3)}{8}(1-x^2)\right]^2 \leq x^n P_n(1/x) \leq \exp(n\sqrt{1-x^2}).$$

For  $|x| \leq 1$  we have  $P_n(x) = 1 + P_n(\xi(x))(x-1)$  where  $\xi(x) \in [-1,1]$ . Due to Markov's inequality we have for  $|\zeta| \leq 1$ ,  $|P_n'(\zeta)| \leq n^2 \max_{|x| \leq 1} |P_n(x)| = n^2$ . Therefore for  $|x| \leq 1$ ,

$$(8) \quad |P_n(x)| \geq 1 - n^2(1-x).$$

We are now ready to estimate  $I_n$  and  $J_n$ . We first show that  $I_n \geq c_1 n^{-2}$  for a positive  $c_1$  which does not depend on  $n$ . By the right inequality of (7) we get

$$\begin{aligned} I_n &= \int_0^1 \left[ \frac{P_n(x)}{x^n P_n(1/x)} \right]^2 dx \geq 2 \int_0^1 |P_n(x)|^2 \exp(-2n\sqrt{1-x^2}) dx \\ &\geq 2 \int_a^1 |P_n(x)|^2 \exp(-2n\sqrt{1-x^2}) dx =: \tilde{I}_n, \\ &\quad \text{where } a = 1 - 1/(2n^2). \end{aligned}$$

Since  $\exp(-2n\sqrt{1-x^2})$  monotonically increases, due to (8) we have

$$\begin{aligned} \tilde{I}_n &\geq 2 \exp(-2n\sqrt{1-a^2}) \int_a^1 (1 - n^2(1-x))^2 dx \\ &\geq 2 \exp(-2) \frac{1}{3n^2} [1 - (1 - n^2(1-a))^3] = \frac{7 \exp(-2)}{12 n^2}. \end{aligned}$$

Thus 
$$I_n \geq \frac{7}{12} \exp(-2) n^{-2}.$$

To complete the proof it is enough to show that  $J_n \leq c_2 n^{-2}$  where  $c_2$  does not depend on  $n$ . By the left inequality of (7),

for  $w = (n-2)(n-3)/8$  we get

$$\begin{aligned}
 J_n &= 2 \int_0^1 \left| \frac{P_n(x)}{x^n P_n(1/x)} \right| dx \leq 2 \int_0^1 \frac{dx}{[1+w(1-x^2)]^2} \\
 &= 2(1+w)^{-2} \int_0^1 \frac{dx}{\left[1 - \frac{w}{1+w} x^2\right]^2} \leq 2(1+w)^{-2} \int_0^1 \frac{dx}{\left[1 - \frac{w}{1+w} x\right]^2} \\
 &= 2(1+w)^{-1} = \frac{16}{n^2 - 5n + 14}.
 \end{aligned}$$

This gives the desired inequality and completes the proof.

To understand the bad properties of Gauss information, suppose one needs to find an  $\epsilon$ -approximation, i.e., to compute  $I = I(f)$  such that  $|\int_0^1 f(x)dx - I(f)| \leq \epsilon$  for all  $f$  from  $F$ . To get  $I(f)$  we use  $n$  samples of  $f$ . From (2) and (4) we conclude that the minimal number of samples  $n$  has to be about  $\ln^2 1/\epsilon$ . Using the information  $N_n^*$  of (3) with  $n = \lceil 0.08\pi^{-2} \ln^2 1/\epsilon \rceil$ , the quadrature  $Q_n^*(N_n^*(f))$  of (3) yields an  $\epsilon$ -approximation since  $e(N_n^*, Q_n^*) \leq \epsilon$ . The cost of  $Q_n^*$  is proportional to  $n$ . From this we conclude that the  $\epsilon$ -complexity, i.e., the minimal cost of computing an  $\epsilon$ -approximation is given by

$$\text{comp}(\epsilon) = \Theta(\ln^2 1/\epsilon)$$

and  $Q_n^*$  is an almost optimal complexity algorithm.

Suppose now one wants to find an  $\epsilon$ -approximation using  $n$  Gauss nodes. Then, due to Theorem 1,  $n$  has to be of order  $\epsilon^{-1/2}$  and the  $\epsilon$ -complexity (minimal cost) of Gauss information is given by

$$\text{comp}_G(\epsilon) = \Theta(\epsilon^{-1/2}).$$

Let  $\text{pen}(\epsilon) := \text{comp}_G(\epsilon)/\text{comp}(\epsilon)$  be the penalty of using Gauss information instead of the optimal one. From this we get

Theorem 2:

$$\text{pen}(\epsilon) = \Theta(\epsilon^{-1/2} \ln^{-2} 1/\epsilon),$$

so that

$$\lim_{\epsilon \rightarrow 0} \text{pen}(\epsilon) = +\infty.$$

## 2.2 Asymptotic setting.

In the worst-case setting the error of an algorithm is defined for fixed information  $N_n$  and for the worst integrand  $f$ . In some situations we prefer to fix  $f$  and apply to it information  $N_n$  with  $n$  tending to infinity. This is called the asymptotic setting. In this setting, information is an infinite sequence

$$\bar{N}(f) = [f(x_1), f(x_2), \dots, f(x_k), \dots].$$

We stress that the points  $x_i$  can be chosen adaptively, i.e.,

$x_i$  can be an arbitrary function of  $f(x_1), f(x_2), \dots, f(x_{i-1})$ .

By an algorithm  $\bar{\varphi}$  using  $\bar{N}$  we now mean a sequence

$\bar{\varphi} = \{\varphi_n\}_{n=1}^{\infty}$  where  $\varphi_n$  uses  $N_n(f) = [f(x_1), \dots, f(x_n)]$ ,

i.e.,  $\varphi_n: N_n(F) \rightarrow \mathbb{R}$ . The  $n$ -th error of  $\bar{\varphi}$  at  $f$  is defined as

$$e_n(\bar{\varphi}, f) := \left| \int_{-1}^1 f(x) dx - \varphi_n(N_n(f)) \right|.$$

In the asymptotic setting we wish to choose an algorithm

$\bar{\varphi}$  as well as the nodes  $x_k$  for which the sequence  $e_n(\bar{\varphi}, f)$  goes to zero as fast as possible for all  $f$  from  $F$ .

Recently, Trojan [15] showed a surprising relation between the worst-case and asymptotic settings. For the integration problem his results can be summarized as follows. (The quantities  $e(N_n)$  and  $e(n)$  are defined as in Section 2.1.)

Given  $f \in F$ , let  $N_n^f$  denote the following nonadaptive information

$$N_n^f(g) = [g(\hat{x}_1), g(\hat{x}_2), \dots, g(\hat{x}_n)],$$

where  $\hat{x}_1 = x_1$  and  $\hat{x}_i = \hat{x}_i(f(\hat{x}_1), \dots, f(\hat{x}_{i-1}))$  for  $i = 2, 3, \dots, n$ .

(i) For any information  $\bar{N}$ , any algorithm  $\bar{\varphi}$  using  $\bar{N}$  and any nonnegative sequence  $\{\delta_n\}_{n=1}^{\infty}$  converging to zero, the set  $F_0$  of  $f$  for which

$$e_n(\bar{\varphi}, f) = o(\delta_n e(N_n^f))$$

is boundary.

(ii) There exist information  $\bar{N}^*$  and an algorithm  $\bar{\varphi}^*$  using  $\bar{N}^*$  such that

$$e_n(\bar{\varphi}^*, f) \leq e(\lfloor n/4 \rfloor).$$

Remark: The statement in (i) that  $F_0$  is boundary means that  $\overline{F - F_0} = F$ . That is, for any nonnegative sequence  $\{\delta_n\}_{n=1}^{\infty}$  converging to zero, the set of  $f$  for which

$$\limsup_{n \rightarrow \infty} \frac{e_n(\bar{\varphi}, f)}{\delta_n e(N_n^f)} > 0$$

is dense in  $F$ .

From (i) with  $\delta_n = \exp(-\sqrt{n})$  and from (2) it follows that for arbitrary information  $\bar{N}$  and an arbitrary algorithm  $\bar{\varphi}$  using  $\bar{N}$  the set of  $f$  for which

$$(9) \quad e_n(\bar{\varphi}, f) = o(\exp(-(1 + \frac{5\pi}{\sqrt{2}})\sqrt{n}))$$

is boundary.

Let  $N_n^*$  and  $Q_n^*$  be given by (3). Define information

$$(10) \quad \bar{N}^*(f) := [N_1^*(f), N_2^*(f), N_4^*(f), \dots, N_{2^k}^*(f), \dots]$$

and the algorithm  $\bar{\varphi}^* = \{\varphi_n^*\}$  using  $\bar{N}^*$  as

$$(11) \quad \varphi_n^*(N_n^*(f)) := Q_{2^{k-1}}^*(N_{2^{k-1}}^*(f)), \quad k = \lfloor \log_2(n+1) \rfloor,$$

where  $N_n^*$  consists of the first  $n$  samples of  $\bar{N}^*$ . From (4)

we get

$$(12) \quad -e_n(\bar{\varphi}^*, f) \leq \exp(-\pi\sqrt{(n+1)/8}),$$

for all  $f$  from  $F$ . Due to (10),  $\bar{N}^*$  and  $\bar{\varphi}^*$  are almost optimal in the asymptotic setting.

We now compare  $\bar{N}^*$  to corresponding Gauss information.

$$(13) \quad \bar{N}^G(f) := [N_1^G(f), N_2^G(f), N_4^G(f), \dots, N_{2^k}^G(f), \dots]$$

where  $N_{2^k}^G$  is given by (5). We prove that Gauss information is also very poor in the asymptotic setting.

Theorem 3: For any algorithm  $\bar{\varphi}$  using  $\bar{N}^G$  and any nonnegative sequence  $\{\delta_n\}$  converging to zero the set of  $f$  for which  $e_n(\bar{\varphi}, f) = o(n^{-2}\delta_n)$  is boundary.

Proof: For each positive integer  $n$ , let  $k = \lfloor \log_2(n+1) \rfloor$  and  $\tilde{N}_n^G(f) := [N_1^G(f), N_2^G(f), \dots, N_{2^{k-1}}^G(f)]$ . We first estimate  $e(\tilde{N}_n^G)$ . The same arguments as in the proof of (6) with

$$h(z) = \prod_{j=0}^{k-1} \left[ \frac{P_{2^j}(z)}{z^{2^j} P_{2^j}(1/z)} \right]^2$$

lead to the inequality

$$(14) \quad e(\tilde{N}_n^G) \geq \int_{-1}^1 h^2(x) dx =: I_k.$$

Let  $b := \sqrt{1 - 2^{-2k+1}}$ . Due to the right hand side of (7) for  $x \in [b, 1]$  we have

$$\left[ \prod_{j=0}^{k-1} x^{2^j} P_{2^j}(1/x) \right]^2 \leq \exp(2 \sum_{j=0}^{k-1} 2^j \sqrt{1-b^2}) \leq \exp(2\sqrt{2}).$$

From (8) for  $x \in [b, 1]$  we get

$$\begin{aligned} \prod_{j=0}^{k-1} P_{2^j}^2(x) &\geq \prod_{j=0}^{k-1} [1 - 2^{2j} (1 - \sqrt{1-2^{-2k+1}})]^2 \geq \prod_{j=0}^{k-1} (1-2^{2j-2k+1})^2 \\ &= \prod_{j=1}^{k-1} (1 - 2^{-2j-1})^2 \geq \prod_{j=0}^{\infty} (1 - 2^{-2j-1})^2 =: c. \end{aligned}$$

Observe that  $c$  exists and is positive. Hence  $h^2(x) \geq c \exp(-2\sqrt{2})$  for  $b \leq x \leq 1$ . We now estimate  $I_k$ .

$$\begin{aligned} I_k &= 2 \int_0^1 h^2(x) dx \geq 2 \int_b^1 h^2(x) dx \geq 2(1-b)c \exp(-2\sqrt{2}) \\ &= 2(1 - \sqrt{1-2^{-2k+1}})c \exp(-2\sqrt{2}) \geq 2^{-2k+1}c \exp(-2\sqrt{2}). \end{aligned}$$

Since  $2^{-2k} \geq n^{-2}/4$ , (14) implies

$$e(\tilde{N}_n^G) \geq c_1 n^{-2}$$

for a positive  $c_1$  which does not depend on  $n$ . Note that this estimate is sharp. Indeed, Theorem 1 yields  $e(N_{2^{k-1}}^G) = \Theta(n^{-2})$ . Obviously  $e(N_{2^{k-1}}^G) \geq e(\tilde{N}_n^G)$ . Therefore

$$e(\tilde{N}_n^G) = \Theta(n^{-2}).$$

This and (i) complete the proof.

Theorem 3 states that the speed of convergence of algorithms using Gauss information is at most  $n^{-2}$  whereas (9) and (12) state that the optimal speed is roughly  $\exp(-c\sqrt{n})$  where  $c > 0$ .



We now show the superiority of the optimal algorithm. Let  $s_1(n) = n^{-2}$  and  $s_2(n) = \exp(-c\sqrt{n})$ . Assume that one wants to choose the minimal  $n$  such that  $s_i(n) \leq \epsilon$ . Then for the function  $s_1$  we have  $n = n_1(\epsilon) = \Theta(\epsilon^{-1/2})$  whereas for the function  $s_2$  we have  $n = n_2(\epsilon) = \Theta(\ln^2 1/\epsilon)$ . The penalty function

$$\text{pen}(\epsilon) := \frac{n_1(\epsilon)}{n_2(\epsilon)} = \Theta(\epsilon^{-1/2} \ln^{-2} 1/\epsilon)$$

goes to  $+\infty$  as  $\epsilon$  goes to zero.

### 3. Gauss-quadrature is almost optimal when $r > 1$ .

In this section we show that Gauss quadrature is almost optimal for the class  $F(D_r)$  with  $r > 1$  in both the worst-case and asymptotic settings.

We begin with the worst-case setting. The quantities  $e(N_n, \varphi)$ ,  $e(N_n)$ , and  $e(n)$  are defined as in Section 2.1, except that now  $F = F(D_r)$ . To show Gauss quadrature is almost optimal, we need some auxiliary results for the integration problem for a different class of integrands. Let  $E_q$  be an ellipse whose foci are  $\pm 1$  and sum of semi-axes is  $q > 1$ . By  $F(E_q)$  we mean the set of functions  $f: [-1, 1] \rightarrow \mathbb{R}$  having an analytic extension  $\bar{f}$  to  $E_q$  such that  $|\bar{f}| \leq 1$  on  $\bar{E}_q$ . For the class  $F(E_q)$  Bakhvalov [1] proved (see also [6], [7], [8], [12]) that the minimal worst-case error of algorithms using  $n$  samples of  $f$  is  $\Theta(q^{-2n})$ . Furthermore he showed that the worst-case error of Gauss quadrature is of order  $q^{-2n}$ . Thus Gauss quadrature is almost optimal in the worst-case setting for the class  $F(E_q)$ .

We shall use Bakhvalov's results to show that Gauss quadrature is also almost optimal for  $F(D_r)$  with  $r > 1$ . Let  $q_1 = r + \sqrt{r^2 + 1}$  and  $q_2 = r + \sqrt{r^2 - 1}$ . Since  $E_{q_2} \subset D_r \subset E_{q_1}$ , we have  $F(E_{q_1}) \subset F(D_r) \subset F(E_{q_2})$ . Therefore there exist positive constants  $c_1$  and  $c_2$ , independent of  $n$ , such that

$$(15) \quad c_1 q_1^{-2n} \leq e(n) \leq e(N_n^G, G_n) \leq c_2 q_2^{-2n}.$$

Here  $N_n^G$ ,  $G_n$  are Gauss information and Gauss quadrature respectively. Observe that  $q_2/q_1 < 1$  and for large  $r$  the ratio  $q_2/q_1$  is close to one. Since  $r > 1$ ,  $(q_2/q_1)^{-2n}$  goes to zero as  $n$  tends to infinity. Thus there is a large gap in the bounds in the estimate (15). In contrast to this, the  $\epsilon$ -complexity is known to within a constant. (The  $\epsilon$ -complexity is defined as at the end of Section 2.1, except that now  $F = F(D_r)$ .) Since  $\text{comp}(\epsilon) = \Theta(\inf\{n: e(n) \leq \epsilon\})$ , (15) yields

**Theorem 4:** For the class  $F(D_r)$  with  $r > 1$  the  $\epsilon$ -complexity of the integration problem in the worst-case setting is

$$\text{comp}(\epsilon) = \Theta(\ln 1/\epsilon).$$

Furthermore, Gauss quadrature  $G_n$  with  $n = \lceil \ln(c_2/\epsilon)/(2 \ln q_2) \rceil$  yields an  $\epsilon$ -approximation with almost minimal complexity.

We now turn to the asymptotic setting. From Trojan's result (i) with  $F = F(D_r)$  and from (15) we get

**Theorem 5:** For arbitrary information  $\bar{N}$ , any algorithm  $\bar{c}$  using  $\bar{N}$  and any nonnegative sequence  $\{\delta_n\}$  converging to zero, the set of  $f$  for which

$$e_n(\bar{\varphi}, f) = o(\delta_n q_1^{-2n})$$

is boundary.

Furthermore, for the algorithm  $\bar{\varphi}^G = \{\varphi_n^G\}$  defined as the following sequence of Gauss quadratures

$$\varphi_n^G(N_n(f)) := G_{2^{k-1}}(N_{2^{k-1}}(f)), \quad k = \lfloor \log_2(n+1) \rfloor,$$

where  $N_n$  consists of  $n$  samples of  $\bar{N}^G$  from (12), we have

$$e_n(\bar{\varphi}^G, f) = O(q_2^{-2n}), \quad \forall f \in F(D_r).$$

Theorem 5 states that the speed of convergence of Gauss quadratures is at least  $q_2^{-2n}$  whereas the speed of convergence of any algorithm is at most  $q_1^{-2n}$ . We stress that there is no practical difference between the functions  $q_1^{-2n}$  and  $q_2^{-2n}$ . Indeed, suppose we choose the minimal  $n$  such that  $q_i^{-2n} \leq \epsilon$ . Then

$$n = n_i(\epsilon) = \left\lceil \frac{\ln 1/\epsilon}{2 \ln 1/q_i} \right\rceil = \Theta(\ln 1/\epsilon).$$

Since  $n_1(\epsilon)$  is of the same order as  $n_2(\epsilon)$ , we have to perform roughly the same number of function evaluation to make the error smaller than  $\epsilon$ , whether the speed of convergence is  $q_1^{-2n}$  or  $q_2^{-2n}$ . This establishes that Gauss quadrature is almost optimal in the asymptotic setting for the class  $F(D_r)$  with  $r > 1$ .

Acknowledgments.

We are indebted to N.S. Bakhvalov for suggesting how to find a lower bound on  $I_n$  of (6). We would also like to thank J.F. Traub for his valuable comments on an earlier version of this manuscript.

Bibliography

- [1] Bakhvalov, N.S., "On the optimal speed of integrating analytic functions" (in Russian), Zh. Vychisl. Mat. Mat. Fiz. 7 (1967), 1011-1020 [English transl.: U.S.S.R. Computational Math. and Math. Phys. 7 (1967), 63-75].
- [2] Bakhvalov, N.S., "On the optimality of linear methods for operator approximation in convex classes of functions" (in Russian), Zh. Vychisl. Mat. Mat. Fiz. 11 (1971), 1014-1018 [English Transl.: U.S.S.R. Computational Math. and Math. Phys. 11 (1971), 244-249].
- [3] Barnhill, R.E., "Asymptotic properties of minimum norm and optimal quadratures," Numer. Math. 12 (1968), 384-393.
- [4] Bojanov, B.D., "Best quadrature formula for a certain class of analytic functions," Zastos. Mat. 14 (1974), 441-447.
- [5] Bojanov, B.D., "Optimal rate of integration and  $\epsilon$ -entropy of a class of analytic functions" (in Russian), Mat. Zametki 14 (1973), 3-10 [English Transl.: Math. Notes 19 (1973), 551-556].
- [6] Chawla, M.M., "Asymptotic estimates for the error in the Gauss-Legendre quadrature formula," Comput. J. 11 (1968), 339-340.
- [7] Chawla, M.M., and Jain, M.K., "Error estimates for Gauss quadrature formulas for analytic functions," Math. Comp. 22 (1968), 82-90.
- [8] Chawla, M.M., and Jain, M.K., "Asymptotic error estimates for the Gauss quadrature formula," Math. Comp. 22 (1968) 91-97.
- [9] Gautschi, W. and Varga, R.S., "Error Bounds for Gaussian Quadrature of Analytic Functions", SIAM J. Numer. Anal. 25, 4, (1983).
- [10] Larkin, F.M., "Optimal approximation in Hilbert spaces with reproducing kernel functions", Math. Comp. 24 (1970), 911-921.

- [11] Pinkus, A., "Asymptotic minimum norm quadrature formulae", Numer. Math. 24 (1975), 163-175.
- [12] - Stenger, F., "Bounds on the error of Gauss-type quadratures", Numer. Math. 8 (1966), 150-160.
- [13] Szegő. G., Orthogonal Polynomials, American Mathematical Society, New York, 1939.
- [14] Traub, J.F., and Woźniakowski, H., A General Theory of Optimal Algorithms, Academic Press, New York, 1980.
- [15] Trojan, J.M., "Asymptotic model for linear problems" (in preparation).