JOURNAL OF APPROXIMATION THEORY 67, 199–214 (1991)

Bounds for the Number of Nodes in Chebyshev Type Quadrature Formulas

PATRICK RABAU*

Department of Mathematics, The Ohio State University, Columbus, Ohio 43210, U.S.A.

AND

BELA BAJNOK

Department of Applied Mathematics, The University of Houston-Downtown, Houston, Texas 77002, U.S.A.

Communicated by Paul Nevai

Received August 22, 1990; revised December 3, 1990

We consider Chebyshev type quadrature formulas on an interval, i.e., quadrature formulas where all nodes are weighted equally. Using a topological method, we give an upper bound for the minimum number of nodes needed in order to achieve a certain degree of precision. We also consider the corresponding problem on the d-dimensional sphere S^d.

1. INTRODUCTION

In this paper we are concerned with Chebyshev type quadrature formulas on a finite interval $I = [a, b]$, i.e., quadrature formulas where all nodes are weighted equally:

$$\frac{1}{\mu(I)} \int_I f(x)\, d\mu(x) \sim \frac{1}{n} \sum_{i=1}^{n} f(x_i).$$

Here μ is a suitable measure (as detailed in Sec. 2) and the nodes x_i are required to lie in the interval I. We call the set of nodes a *μ-averaging set* of degree p if the formula is exact for all polynomials of degree up to p. (See [7] for a survey article on such quadrature formulas.)

Several authors have found upper bounds for the degree of precision p

* Partially supported by NSF Grant DMS-8801182-02.

0021-9045/91 $3.00

of such a formula when the number of nodes n is fixed; equivalently, if the degree of precision is fixed, they obtain a lower bound for the number of nodes. Bernstein [3] showed $n > (1/16)p^2$ for the Lebesgue measure. Later authors [8, 5] have considered the case of Jacobi weight functions.

The problem we want to address is to find an *upper bound* for the minimum number of nodes needed in order to achieve a certain degree of precision. Using a topological method, which was inspired by [1], we develop in Section 2 a general formula for such an upper bound (Theorem 2.1). We actually treat the more general situation of μ-averaging sets for an arbitrary finite dimensional vector space of continuous functions, not necessarily polynomials.

This is specialized in Section 3 to the case of polynomials integrated with respect to a Jacobi weight function. The upper bound we are able to compute (Theorem 3.1) differs from the lower bound in [5] only by a factor p. For example, in the case of the Lebesgue measure, we obtain $N(p) = O(p^3)$ for the minimum size $N(p)$ of a μ-averaging set of degree p. (It is probable that the lower bound of Bernstein is closer to the truth.)

Finally, in Section 4 we consider averaging sets on the d-dimensional sphere $S^d = \{(x_0, ..., x_d) \mid x_0^2 + \cdots + x_d^2 = 1\}$ in $\mathbf{R}^{d+1}$. (See the book of Stroud [11] for specific examples.) Combinatorists have developed a more recent interest in these averaging sets, which they call spherical designs (see [6]).

By the very simple observation that for integration purposes the sphere is equivalent to a product of intervals with suitable ultraspherical weight functions, we can construct (Theorem 4.3) spherical averaging sets as "products" of interval averaging sets. Letting $N_d(p)$ denote the minimum size of an averaging set of degree p on S^d, we get the estimate (Theorem 4.1)

$$N_d(p) = O(p^{d^2/2 + 3d/2 - 1}). \tag{1.1}$$

We can actually show the existence of a spherical averaging set of size n and degree p for every $n \geqslant \tilde{N}_d(p)$, where $\tilde{N}_d(p)$ has the same order of magnitude as (1.1). For the sphere S^2, this improves and simplifies previous results of one of the authors [2]. We should emphasize that our methods are elementary and do not require knowledge of spherical harmonics on S^d.

2. The Main Result

Let μ be a Borel measure on the finite interval $I = [a, b]$ which satisfies the following two properties:

$$\begin{aligned}&\text{(i)} \quad \mu(\{x\})=0 \text{ for every } x\in I;\\ &\text{(ii)} \quad \mu(J)>0 \text{ for every subinterval } J \text{ of } I \text{ of positive length.}\end{aligned} \tag{2.1}$$

Condition (ii) ensures that the formula

$$(f, g)\mapsto \int_I fg\,d\mu \tag{2.2}$$

defines a positive definite inner product on the space of real-valued continuous functions on I. Condition (i) is used in the proof of Theorem 2.1 below. The important case for the applications is $d\mu(x)=w(x)\,dx$ for some Lebesgue integrable weight function $w(x)$ taking positive values a.e.

Let V be a finite dimensional vector space of real-valued continuous functions on I. We are interested in Chebyshev type quadrature formulas exact for all functions in V. In other words we consider *μ-averaging sets* for functions in V, i.e., sets of distinct points $\{x_1, ..., x_n\}\subseteq I$ such that

$$\frac{1}{\mu(I)}\int_I f\,d\mu=\frac{1}{n}\sum_{i=1}^{n} f(x_i) \tag{2.3}$$

for all $f\in V$.

In Theorem 2.1 below, we give an upper bound for the minimum size of such an averaging set. We actually exhibit a number N_0 such that for every integer $n>N_0$ there exists a μ-averaging set for V of size n. (The existence of such a number N_0 has been proved under much more general circumstances by Seymour and Zaslavsky [9], but their method does not yield explicit estimates.)

Before stating our main result, let us make a few simplifying assumptions. By the linearity of the problem, it is enough to check (2.3) for functions $f_1, ..., f_m$ forming a basis for V. This in turn is equivalent to the single vector-valued equation

$$\frac{1}{\mu(I)}\int_I \mathbf{f}\,d\mu=\frac{1}{n}\sum_{i=1}^{n} \mathbf{f}(x_i), \tag{2.4}$$

where $\mathbf{f}=(f_1, ..., f_m)$. Also, since (2.3) always holds for constant functions f independently of the choice of the nodes x_i, we may subtract from each function $f\in V$ its μ-average

$$\bar{f}=\frac{1}{\mu(I)}\int_I f\,d\mu$$

and make the following assumption: *any function in V has μ-average zero.*

The problem is then reduced to finding a set of points $x_i \in I$ such that

$$\sum_{i=1}^{n} \mathbf{f}(x_i) = 0.$$

We will use the following general notation. The Euclidean inner product and Euclidean norm in $\mathbf{R}^m$ will be written $\langle u, v \rangle$ and $\|v\|$. For (continuous) functions $f: I \to \mathbf{R}$ and $\mathbf{f}: I \to \mathbf{R}^m$, we write

$$\|f\|_\infty = \max_{x \in I} |f(x)| \qquad \text{and} \qquad \|\mathbf{f}\|_\infty = \max_{x \in I} \|\mathbf{f}(x)\|.$$

The length of an interval J is denoted by $|J|$.

THEOREM 2.1. *Let $\mathbf{f} = (f_1, ..., f_m)$ be a Lipschitz function from I to $\mathbf{R}^m$ with Lipschitz constant*

$$L = \sup \left\{ \frac{\|\mathbf{f}(x) - \mathbf{f}(y)\|}{|x - y|} \,\middle|\, x, y \in I, x \neq y \right\},$$

having μ-average $\bar{\mathbf{f}} = 0$, and such that the components $f_1, ..., f_m$ are linearly independent. Let λ be the smallest eigenvalue of the (positive definite) Gram matrix

$$\left(\int_I f_i f_j \, d\mu \right)_{i,j=1}^{m}. \tag{2.5}$$

Then for every integer $n > N_0$ with

$$N_0 = \frac{|I| \; \mu(I) \; \|\mathbf{f}\|_\infty \; L}{2\lambda}$$

there is a set of n distinct points $\{x_1, ..., x_n\} \subseteq I$ such that

$$\sum_{i=1}^{n} \mathbf{f}(x_i) = 0.$$

Proof. Let B^m and S^{m-1} denote, respectively, the unit ball and the unit sphere in $\mathbf{R}^m$. Fix $\eta > 1$. For each $u \in B^m$, define the continuous function $h_u: I \to \mathbf{R}$ by

$$h_u(x) = \eta \, \|\mathbf{f}\|_\infty + \langle u, \mathbf{f}(x) \rangle.$$

Note that

$$\int_I h_u \, d\mu = \eta \mu(I) \, \|\mathbf{f}\|_\infty$$

(because $\bar{\mathbf{f}}=0$), and by the Cauchy–Schwarz inequality

$$h_u(x) \geqslant (\eta - 1)\, \|\mathbf{f}\|_\infty .$$

Now fix a positive integer n (to be determined later) and define, for each $u \in B^m$, an increasing sequence $x_i = x_i(u)$ $(i = 1, ..., n)$ of points in I as follows. Decompose I into a union of n nonoverlapping intervals $I = I_1 \cup \cdots \cup I_n$ such that, for each i,

$$\int_{I_i} h_u \, d\mu = \frac{1}{n} \int_I h_u \, d\mu = \frac{\eta \mu(I)\, \|\mathbf{f}\|_\infty}{n}, \tag{2.6}$$

and let x_i be the midpoint of I_i. Since h_u is bounded away from zero, conditions (2.1) on μ ensure that the intervals I_i are uniquely determined and that $x_i(u)$ varies continuously with u.

Define the continuous map $\Phi \colon B^m \to \mathbf{R}^m$ by

$$\Phi(u) = \frac{\eta \mu(I)\, \|\mathbf{f}\|_\infty}{n} \sum_{i=1}^{n} \mathbf{f}(x_i(u)).$$

We will show that, for n large enough, $\Phi(u) = 0$ for some $u \in B^m$. The corresponding sequence $x_1(u), ..., x_n(u)$ is then an averaging set for $\mathbf{f}$.

To establish our claim, we introduce the linear map Φ_1 from $\mathbf{R}^m$ to itself defined by

$$\Phi_1(u) = \int_I h_u \mathbf{f} \, d\mu = \int_I \langle u, \mathbf{f} \rangle \mathbf{f} \, d\mu$$

(the second equality holds because $\bar{\mathbf{f}} = 0$), and write $\varphi_1 = \Phi_1 |_{S^{m-1}}$. The (i, j)-entry of the matrix representing Φ_1 with respect to the standard basis $\{e_1, ..., e_m\}$ of $\mathbf{R}^m$ is

$$\langle e_i, \Phi_i(e_j) \rangle = \int_I f_i f_j \, d\mu,$$

i.e., the (i, j)-entry of the Gram matrix (2.5). Since this matrix is nonsingular, φ_1 is homotopically nontrivial as a map into $\mathbf{R}^m \setminus \{0\}$. Moreover,

$$\min_{\|u\| = 1} \|\varphi_1(u)\| = \lambda. \tag{2.7}$$

Writing $\varphi = \Phi |_{S^{m-1}}$, we show that, for n large enough,

$$\|\varphi(u) - \varphi_1(u)\| < \|\varphi_1(u)\| \qquad \text{for all} \quad u \in S^{m-1}, \tag{2.8}$$

which implies that φ has no zero, φ is homotopic to φ_1 as a map into

$\mathbf{R}^m\backslash\{0\}$ (move $\varphi(u)$ towards $\varphi_1(u)$ along the straight segment joining the two points), and hence φ is homotopically nontrivial as a map into $\mathbf{R}^m\backslash\{0\}$. This implies that Φ has a zero in B^m (otherwise the map $\varphi\colon S^{m-1}\to\mathbf{R}\backslash\{0\}$ would factor through the contractible space B^m and hence would be homotopically trivial).

We have

$$\varphi(u)=\sum_{i=1}^{n}\int_{I_i} h_u(x)\,\mathbf{f}(x_i)\,d\mu(x)$$

by (2.6), and

$$\varphi_1(u)=\sum_{i=1}^{n}\int_{I_i} h_u(x)\,\mathbf{f}(x)\,d\mu(x).$$

Moreover, if $x\in I_i$, then

$$\|\mathbf{f}(x_i)-\mathbf{f}(x)\|\leqslant L\,|x_i-x|\leqslant\tfrac{1}{2}\,|I_i|\,L.$$

Therefore,

$$\begin{aligned}\|\varphi(u)-\varphi_1(u)\| &\leqslant\sum_{i=1}^{n}\int_{I_i} h_u(x)\,\|\mathbf{f}(x_i)-\mathbf{f}(x)\|\,d\mu(x)\\ &\leqslant\frac{1}{2}L\sum_{i=1}^{n}|I_i|\int_{I_i} h_u\,d\mu\\ &=\frac{1}{2}L\frac{\eta\mu(I)\,\|\mathbf{f}\|_\infty}{n}\sum_{i=1}^{n}|I_i|\\ &=\frac{\eta\,|I|\,\mu(I)\,\|\mathbf{f}\|_\infty\,L}{2n}.\end{aligned}$$

By (2.7), (2.8) will hold as soon as

$$\frac{\eta\,|I|\,\mu(I)\,\|\mathbf{f}\|_\infty\,L}{2n}<\lambda.$$

Since $\eta>1$ is arbitrary, we deduce the existence of a μ-averaging set for $\mathbf{f}$ with n points whenever

$$n>\frac{|I|\,\mu(I)\,\|f\|_\infty\,L}{2\lambda}.\quad\blacksquare$$

Let us write $N_0=N_0(\mathbf{f})$ to show the dependence on $\mathbf{f}$. Notice that $N_0(\mathbf{f})$ is unchanged if μ is multiplied by a positive constant, or if each f_i is multiplied by a common nonzero constant. Also, if $\mathbf{f}$ is differentiable, then the Lipschitz constant is $L = \|\mathbf{f}'\|_\infty$.

Remarks 2.2. (1) The argument used in the proof can be adapted to the case where the functions f_i are just continuous (one uses the fact that condition (2.1)(ii) on μ implies $|J| \to 0$ as $\mu(J) \to 0$ uniformly for all subintervals J of I). But the result is not as clean and, since it is not needed for polynomials, we will not give any detail.

(2) If the average $\bar{\mathbf{f}}$ is not zero, the result has to be modified as follows: there is an averaging set for $\mathbf{f}$ with n points whenever

$$n > \frac{|I|\, \mu(I)\, \|\mathbf{f} - \bar{\mathbf{f}}\|_\infty\, L}{2\lambda},$$

where λ is computed from the modified Gram matrix

$$\left(\int_I (f_i - \bar{f}_i)(f_j - \bar{f}_j)\, d\mu \right)_{i,j}$$

(assuming that $f_1, ..., f_m$ and the constant 1 are linearly independent).

(3) (Symmetry) Since the whole situation is invariant under a linear change of variable, we may assume without loss of generality that we are working on the interval $[-1, 1]$. Let us assume for simplicity that the measure μ is given by a weight function $w(x)$. If $w(x)$ is even and if the even and odd parts of any function in V also belong to V (i.e., V can be decomposed into a direct sum $V = V^+ \oplus V^-$ with V^+ (resp. V^-) consisting entirely of even (resp. odd) functions), then it is natural to consider averaging sets which are symmetric with respect to the origin. Such averaging sets integrate correctly all odd functions. So the problem is reduced to finding symmetric averaging sets for the functions in V^+, for which it is equivalent to work on the interval $[0, 1]$, by symmetry. When $\dim V^+ < \dim V$, this already gives an improvement of the bound N_0. As we will see in Section 3, for polynomials this can be further improved by the change of variables $t = x^2$, $x \in [0, 1]$, as explained in the next remark.

(4) (Change of variable) For simplicity, let us assume that $\mathbf{f}$ is differentiable and that the measure μ is given by a weight function $w(x)$. If we perform the change of variables $x = \varphi(\theta)$, where φ is a differentiable monotone function from an interval J onto I, the measure $d\mu(x) = w(x)\, dx$ is replaced by $d\tilde{\mu}(\theta) = w(\varphi(\theta))\, \varphi'(\theta)\, d\theta$ and (2.4) becomes

$$\frac{1}{\tilde{\mu}(J)} \int_J (\mathbf{f} \circ \varphi)\, d\tilde{\mu} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{f}(\varphi(\theta_i)).$$

In the bound given by Theorem 2.1 one has to replace $\mathbf{f}$ by $\mathbf{f} \circ \varphi$ and μ by $\tilde{\mu}$. The quantities corresponding to $\mu(I)$, λ, and $\|\mathbf{f}\|_\infty$ remain unchanged; $|I|$ is replaced by $|J|$ and $L = \|\mathbf{f}'\|_\infty$ by $\|(\mathbf{f} \circ \varphi)'\|_\infty$, so that N_0 becomes

$$\frac{|J|\, \mu(I)\, \|\mathbf{f}\|_\infty\, \|(\mathbf{f} \circ \varphi)'\|_\infty}{2\lambda}.$$

If one can choose φ so that $|J| \cdot \|(\mathbf{f} \circ \varphi)'\|_\infty$ is substantially smaller than $|I| \cdot \|\mathbf{f}'\|_\infty$, as we can for polynomials (see Sect. 3), this results in a better bound for the minimum size of a μ-averaging set.

(5) If one makes a change of basis in V by means of a nonsingular matrix A and the basis $\mathbf{f} = (f_1, ..., f_m)$ of V is sent to $A\mathbf{f}$, the following estimate holds,

$$N_0(A\mathbf{f}) \leqslant (\mathrm{cond}(A))^2\, N_0(\mathbf{f}),$$

where $\mathrm{cond}(A) = \|A\| \cdot \|A^{-1}\|$ is the condition number of the matrix A.

One would like to choose a basis $\mathbf{f}$ of V that minimizes $N_0(\mathbf{f})$. The estimate in Remark (5) is not helpful in that respect, because the condition number of a matrix is never less than 1. If, however, the functions $f_1, ..., f_m$ are orthogonal with respect to the inner product (2.2) on V, it is easily checked that $N_0(\mathbf{f})$ is decreased if each f_i is multiplied by a scalar so that all the f_i have the same L^2-norm. Since multiplying all the f_i simultaneously by a common nonzero scalar does not change $N_0(\mathbf{f})$, we may as well assume in that case that $(f_1, ..., f_m)$ is an orthonormal basis for V.

When doing specific computations in the next section, we will choose $\mathbf{f}$ to be an orthonormal basis for V, which gives $\lambda = 1$. The space V will consist of differentiable functions (actually polynomials). To make the computation easier, we will overestimate $\|\mathbf{f}\|_\infty$ by the more easily computed quantity $\sqrt{\sum_{k=1}^m \|f_k\|_\infty^2}$, and similarly for the Lipschitz constant $L = \|\mathbf{f}'\|_\infty$. Note that $\|\mathbf{f}\|_\infty = \sqrt{\sum_{k=1}^m \|f_k\|_\infty^2}$ if the maxima of the f_k all occur at the same point of I, which is often the case for orthogonal polynomials. If we combine this with Remark 2.2(4) about changes of variable, we get:

COROLLARY 2.3. *Suppose V consists of differentiable functions (all of μ-average zero). Let $(f_1, ..., f_m)$ be an orthonormal basis for V and let $\varphi\colon J \to I$ be a change of variable as in Remark* 2.2(4). *Then there exists a μ-averaging set for V of size n for every integer $n > N_1$ with*

$$N_1 = \frac{1}{2}\, |J|\, \mu(I) \sqrt{\sum_{k=1}^{m} \|f_k\|_\infty^2}\, \sqrt{\sum_{k=1}^{m} \|(f_k \circ \varphi)'\|_\infty^2}.$$

3. Jacobi Weight Functions

In this section we look at averaging sets of degree p with respect to the Jacobi measure $d\mu(x)=(1-x)^{\alpha}(1+x)^{\beta}\,dx$ $(\alpha, \beta>-1)$ on the interval $[-1, 1]$. Let $N(p)$ denote the minimum size of such an averaging set of degree p, and write $\rho=\max(\alpha, \beta)$.

The following lower bounds for $N(p)$ are known,

$$\tfrac{1}{2}p<N(p) \tag{3.1}$$

and

$$C_1 p^{2\rho+2}\leqslant N(p) \qquad \text{if} \quad \rho\geqslant -1/2, \tag{3.2}$$

where C_1 is a constant depending on α, β and independent of p. Formula (3.1) is a standard fact, valid for any kind of quadrature formula; (3.2) is proved in [5].

THEOREM 3.1. *With the notation above, we have*

$$N(p)\leqslant C_2 p^2 \qquad \textit{if} \quad \rho<-1/2;$$

$$N(p)\leqslant C_2 p^{2\rho+3} \qquad \textit{if} \quad \rho\geqslant -1/2.$$

Here C_2 is a constant depending on α, β and independent of p. If $\rho\geqslant -1/2$, one can take "asymptotically" (as $p\to\infty$)

$$C_2=\frac{\pi}{2}\frac{\Gamma(\alpha+1)\,\Gamma(\beta+1)}{[\Gamma(\rho+1)]^2\,\Gamma(\alpha+\beta+2)}\frac{1}{\sqrt{(\rho+1)(\rho+2)}}.$$

The word "asymptotically" in the theorem means

$$\limsup_{p\to\infty}\frac{N(p)}{p^{2\rho+3}}\leqslant C_2.$$

Based on the case of Chebyshev polynomials ($\rho=-1/2$, $N(p)\sim p/2$), it is most likely that the exponent $2\rho+3$ is not the best possible in general. (For that reason, there is no point in doing the computation for C_2 more precisely than asymptotically.) One may guess that the best exponent is in fact given by the lower bound:

Conjecture 3.2. $N(p)=O(p^{2\rho+2})$ when $\rho\geqslant -1/2$.

Remark 3.3. In the special case $\alpha\geqslant -1/2$ and $\beta=-1/2$, we have

$$C_2=\frac{\pi 2^{2\alpha}}{\Gamma(2\alpha+2)\sqrt{(\alpha+1)(\alpha+2)}}$$

(use the duplication formula for the gamma function).

Proof of Theorem 3.1. We will use Corollary 2.3, taking f_k ($k=1, ..., p$) to be orthogonal polynomials with respect to the Jacobi weight $w(x)=(1-x)^\alpha (1+x)^\beta$, with each f_k normalized to have L^2-norm equal to 1. We also use the change of variable $x=\varphi(\theta)=\cos\theta$, $\theta\in J=[0,\pi]$, $x\in I=[-1,1]$. So $|J|=\pi$ and

$$\mu(I)=\int_{-1}^{1} w(x)\,dx=\frac{2^{\alpha+\beta+1}\Gamma(\alpha+1)\,\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}.$$

We have $f_k=\alpha_k^{-1/2}P_k^{(\alpha,\beta)}$ where $P_k^{(\alpha,\beta)}$ is the Jacobi polynomial of degree k as in [12] and

$$\begin{aligned}\alpha_k&=\int_{-1}^{1}[P_k^{(\alpha,\beta)}(x)]^2\,w(x)\,dx\\ &=\frac{2^{\alpha+\beta+1}\Gamma(k+\alpha+1)\,\Gamma(k+\beta+1)}{(2k+\alpha+\beta+1)k!\,\Gamma(k+\alpha+\beta+1)}\sim 2^{\alpha+\beta}k^{-1}.\end{aligned}$$

First assume that $\rho\geqslant -1/2$. Then

$$\|P_k^{(\alpha,\beta)}\|_\infty=\binom{k+\rho}{k}\sim\frac{k^\rho}{\Gamma(\rho+1)},$$

which implies

$$\begin{aligned}\sqrt{\sum_{k=1}^{p}\|f_k\|_\infty^2}&=\sqrt{\sum_{k=1}^{p}\alpha_k^{-1}\,\|P_k^{(\alpha,\beta)}\|_\infty^2}\\ &\sim\sqrt{\sum_{k=1}^{p}\frac{k^{2\rho+1}}{2^{\alpha+\beta}[\Gamma(\rho+1)]^2}}\\ &\sim\frac{p^{\rho+1}}{2^{(\alpha+\beta)/2}\Gamma(\rho+1)\sqrt{2\rho+2}}.\end{aligned}$$

To estimate $\|(f_k\circ\varphi)'\|_\infty$, we use an inequality of Bernstein (see, for example, [4]) which asserts that

$$\max_{0\leqslant\theta\leqslant\pi}\left|\frac{d}{d\theta}(P(\cos\theta))\right|\leqslant k\max_{|x|\leqslant 1}|P(x)|$$

for every polynomial P of degree k. This gives

$$\|(f_k\circ\varphi)'\|_\infty\leqslant k\,\|f_k\|_\infty=k\alpha_k^{-1/2}\,\|P_k^{(\alpha,\beta)}\|_\infty$$

and

$$\sqrt{\sum_{k=1}^{p} \|(f_k \circ \varphi)'\|_\infty^2} \leqslant \sqrt{\sum_{k=1}^{p} k^2 \alpha_k^{-1} \|P_k^{(\alpha,\beta)}\|_\infty^2}$$

$$\sim \sqrt{\sum_{k=1}^{p} \frac{k^{2\rho+3}}{2^{\alpha+\beta}[\Gamma(\rho+1)]^2}}$$

$$\sim \frac{p^{\rho+2}}{2^{(\alpha+\beta)/2}\Gamma(\rho+1)\sqrt{2\rho+4}}.$$

Substituting the various estimates above in Corollary 2.3 yields "asymptotically"

$$N(p) \leqslant \frac{\pi}{2} \frac{\Gamma(\alpha+1)\,\Gamma(\beta+1)}{[\Gamma(\gamma+1)]^2\,\Gamma(\alpha+\beta+2)} \frac{1}{\sqrt{(\rho+1)(\rho+2)}}\, p^{2\rho+3}.$$

In the case $\rho < -1/2$, we can only claim $\|P_k^{(\alpha,\beta)}\|_\infty = O(k^{-1/2})$, which gives, by a similar computation, $N(p) = O(p^2)$. ∎

Assume we are now in the ultraspherical case with $\alpha = \beta \geqslant -1/2$. By (3.2) and Theorem 3.1 we know that

$$C_1\, p^{2\alpha+2} \leqslant N(p) \leqslant C_2\, p^{2\alpha+3}.$$

Costabile [5] gives an explicit expression for a constant C_1 in terms of gamma and Bessel functions. For the upper bound, the constant C_2 in Theorem 3.1 can be replaced by $\frac{1}{2}C_2$, as proved below using the symmetry of the weight function $w(x) = (1-x^2)^\alpha$.

PROPOSITION 3.4. *Keeping the notation of Theorem* 3.1, *let* $\alpha = \beta \geqslant -1/2$. *Asymptotically* (*as* $p \to \infty$), *we have*

$$N(p) \leqslant C_2'\, p^{2\alpha+3}$$

with

$$C_2' = \frac{\pi}{4} \frac{1}{\Gamma(2\alpha+2)\sqrt{(\alpha+1)(\alpha+2)}}.$$

Proof. Use the ideas in Remark 2.2(3), which are applicable since $w(x) = (1-x^2)^\alpha$ is even. If we look for symmetric averaging sets, it is enough to look for averaging sets with respect to $(1-x^2)^\alpha\, dx$ on [0, 1] that integrate correctly all even polynomials of degree at most p. By the change of variable $x = t^{1/2}$, $t \in [0, 1]$, the problem is changed into integrating arbitrary polynomials of degree not exceeding $p/2$ with respect to the measure $\frac{1}{2}(1-t)^{\alpha} t^{-1/2}\, dt$, which, by a linear change of variable, is equivalent to the Jacobi weight $(1-x)^{\alpha}(1+x)^{-1/2}$ on $[-1, 1]$. So, in the constant C_2 of Theorem 3.1, take $\rho = \alpha$, $\beta = -1/2$, and replace p by $p/2$, not forgetting to multiply the whole quantity by 2 to take the negative half of $[-1, 1]$ into account. From Remark 3.3 we get (asymptotically) $N(p) \leqslant C_2' p^{2\alpha+3}$ with C_2' as claimed. ∎

Remark 3.5. The case $w(x) = 1$ (that is, $\alpha = \beta = 0$) is worth mentioning explicity,

$$\frac{1}{16} p^2 < N(p) \leqslant \frac{\pi}{4\sqrt{2}} p^3,$$

where the second inequality is only asymptotic $(p \to \infty)$. The first inequality was proved by Bernstein [3] in the form $p < 4\sqrt{N}$.

4. Averaging Sets on the Sphere

In this section, we investigate *averaging sets* of degree p on the d-dimensional unit sphere $S^d = \{x = (x_0, ..., x_d) \in \mathbf{R}^{d+1} \mid x_0^2 + \cdots + x_d^2 = 1\}$, i.e., finite subsets X of S^d such that

$$\frac{1}{\sigma_d(S^d)} \int_{S^d} f(x)\, d\sigma_d(x) = \frac{1}{|X|} \sum_{x \in X} f(x) \tag{4.1}$$

for all polynomials $f(x) = f(x_0, ..., x_d)$ of (total) degree at most p. Here σ_d is the surface measure on S^d.

Delsarte *et al.* [6] give a lower bound for the size of an averaging set X of degree p on S^d:

$$|X| \geqslant \binom{d + \lfloor p/2 \rfloor}{\lfloor p/2 \rfloor} + \binom{d + \lfloor p/2 \rfloor - 1}{\lfloor p/2 \rfloor}$$
$$\sim \frac{1}{2^d d!} p^d \qquad \text{as} \quad p \to \infty.$$

Theorem 4.1. *Let $N_d(p)$ denote the minimum size of an averaging set of degree p on S^d. There is a positive constant $C(d)$ independent of p such that*

$$N_d(p) \leqslant C(d)\, p^{d^2/2 + 3d/2 - 1}.$$

Asymptotically (*as* $p \to \infty$) *one can take*

$$C(d) = \left(\frac{\pi}{2}\right)^{d-1} \sqrt{\frac{6}{(d+1)(d+2)}} \prod_{i=1}^{d} \frac{1}{i!}.$$

Remark 4.2. For the usual sphere S^2, this gives asymptotically

$$N_2(p) \leqslant \frac{\pi}{4\sqrt{2}} p^4.$$

Our method is recursive and based on the simple fact that, for integration purposes, the sphere S^d is equivalent to the product of the "equator" S^{d-1} and of the interval $[-1, 1]$ with a suitable ultraspherical weight function. (Separation of variables occurs when integrating monomials.) Thus one can obtain an averaging set of degree p on S^d by taking suitably spaced $(d-1)$-dimensional sections of S^d parallel to the equator S^{d-1}, and putting an averaging set of degree p on each section.

Before stating the precise result, let us introduce some notation. For $x = (x_0, ..., x_d) \in S^d$, let $\tilde{x} = (x_0, ..., x_{d-1})$ and, for simplicity of notation, write $x_d = t \in [-1, 1]$. We have $\tilde{x} = \sqrt{1-t^2}\, y$ for some $y \in S^{d-1}$, so that $x = (\sqrt{1-t^2}\, y, t)$.

THEOREM 4.3. *Let $Y \subseteq S^{d-1}$ be an averaging set of degree p, and let $T \subseteq (-1, 1)$ be an averaging set of degree p on $[-1, 1]$ with respect to the measure $(1-t^2)^{(d-2)/2}\, dt$. Then*

$$X = \{(\sqrt{1-t^2}\, y, t) \mid y \in Y, t \in T\}$$

is an averaging set of degree p on S^d.

Proof. We will use multi-index notation: if $\alpha = (\alpha_0, ..., \alpha_d)$ is a $(d+1)$-tuple of nonnegative integers, write $x^\alpha = x_0^{\alpha_0} \cdots x_d^{\alpha_d}$; also, let $|\alpha| = \alpha_0 + \cdots + \alpha_d$, $\tilde{\alpha} = (\alpha_0, ..., \alpha_{d-1})$, and $|\tilde{\alpha}| = \alpha_0 + \cdots + \alpha_{d-1}$.

By the linearity of the problem, it is enough to verify (4.1) for all the monomials $f(x) = x^\alpha$ with $|\alpha| \leqslant p$. We have

$$x^\alpha = y^{\tilde{\alpha}} (1-t^2)^{|\tilde{\alpha}|/2}\, t^{\alpha_d}$$

and

$$d\sigma_d(x) = d\sigma_{d-1}(y)(1-t^2)^{(d-2)/2}\, dt.$$

The integral in (4.1) becomes

$$\left(\frac{1}{\sigma_{d-1}(S^{d-1})} \int_{S^{d-1}} y^{\tilde{\alpha}}\, d\sigma_{d-1}(y)\right)\left(K \int_{-1}^{1} (1-t^2)^{|\tilde{\alpha}|/2}\, t^{\alpha_d} (1-t^2)^{(d-2)/2}\, dt\right), \tag{4.2}$$

where K is a constant chosen to normalize the measure. Since $|X| = |Y| \cdot |T|$, the sum in (4.1) becomes

$$\left(\frac{1}{|Y|}\sum_{y\in Y} y^{\tilde{\alpha}}\right)\left(\frac{1}{|T|}\sum_{t\in T}(1-t^2)^{|\tilde{\alpha}|/2}\, t^{\alpha_d}\right). \tag{4.3}$$

By definition of averaging set, the factors of (4.2) and (4.3) involving y coincide whenever $|\tilde{\alpha}| \leqslant p$. If $|\tilde{\alpha}|$ is odd, then these factors are zero and $(4.2) = (4.3)$. If $|\tilde{\alpha}|$ is even, then $(1-t^2)^{|\tilde{\alpha}|/2}\, t^{\alpha_d}$ is a *polynomial* in t of degree at most p, so that the factors of (4.2) and (4.3) involving t also coincide, and hence $(4.2) = (4.3)$. So X is an averaging set of degree p on S^d. ∎

Proof of Theorem 4.1. We apply the recursive procedure of Theorem 4.3. To start the induction, we use the well-known fact that a regular polygon with $p+1$ vertices is an averaging set (of minimum size) of degree p on the circle S^1. Let $N^{(i)}(p)$ denote the minimum size of an averaging set of degree p on $[-1, 1]$ with respect to the ultraspherical measure $(1-t^2)^{(i-2)/2}\, dt$. A repeated application of Theorem 4.3 shows that the minimum size $N_d(p)$ of an averaging set of degree p on S^d satisfies

$$N_d(p) \leqslant (p+1) \prod_{i=2}^{d} N^{(i)}(p). \tag{4.4}$$

By Proposition 3.4, we have

$$N^{(i)}(p) \leqslant C_i\, p^{i+1}, \tag{4.5}$$

where asymptotically (as $p \to \infty$) one can take

$$C_i = \frac{\pi}{2(i-1)!\,\sqrt{i(i+2)}}.$$

A short computation gives the result of the theorem after combining (4.4) and (4.5). ∎

Note that our method depends crucially on the ability to find averaging sets on $[-1, 1]$ with respect to ultraspherical weight functions. Any improvement in the estimates for averaging sets on intervals will automatically yield an improvement for the sphere. On the basis of Conjecture 3.2, we can make a corresponding conjecture for the sphere:

Conjecture 4.4. The minimum size of an averaging set of degree p on S^d is $O(p^{d(d+1)/2})$ as $p \to \infty$.

We conclude the section by examining a quantity related to $N_d(p)$. It is a special case of a result of [9] that *every* sufficiently large integer can occur as the size of an averaging set of degree p on S^d. The next result gives an estimate for the smallest integer $\tilde{N}_d(p)$ such that for every $n \geqslant \tilde{N}_d(p)$ there exists an averaging set of degree p on S^d having size n. The upper bound we give for $\tilde{N}_d(p)$ has the same order of magnitude as the upper bound for $N_d(p)$ in Theorem 4.1.

PROPOSITION 4.5. *With the notation above,* $\tilde{N}_d(p) = O(p^{d^2/2+3d/2-1})$ *as* $p \to \infty$.

The following lemma is not sharp, but is sufficient for our purposes.

LEMMA 4.6. *Let r be a positive integer and let $a_1, ..., a_r$ be real numbers not smaller than 6. Then every integer $n \geqslant 8^{r-1}a_1 \cdots a_r$ can be expressed as a sum of 2^{r-1} terms, each term being a product of the form $n_1 n_2 \cdots n_r$ for integers $n_i \geqslant a_i$ $(i = 1, ..., r)$.*

Proof. The case $r = 1$ is clear. For $r = 2$, assume $a_1 \leqslant a_2$. Since $a_1 \geqslant 6$, there exist two distinct primes p, q between a_1 and $2a_1$ [10, Th. 7, p. 144]. Since p and q are relatively prime, any integer n can be written in the form $n = px + qy$ for some integers x, y. We want to determine a condition on n such that both x and y can be taken $\geqslant a_2$. The number x can be adjusted so that $a_2 \leqslant x \leqslant a_2 + q$. Then $y \geqslant a_2$ as soon as

$$n \geqslant p(a_2 + q) + qa_2. \tag{4.6}$$

Since

$$\begin{aligned} p(a_2+q)+qa_2 &\leqslant 2a_1(a_2+2a_1)+2a_1a_2 \\ &= 4a_1a_2 + 4a_1^2 \leqslant 8a_1a_2, \end{aligned}$$

(4.6) will hold if $n \geqslant 8a_1a_2$.

The general case follows by induction. For simplicity we just show the case $r = 3$. Suppose $n \geqslant 8^2 a_1 a_2 a_3 = 8(8a_1a_2)\, a_3$. By the case $r = 2$, n can be written as a sum of 2 terms, each of the form mn_3 with $m \geqslant 8a_1a_2$ and $n_3 \geqslant a_3$. Each m in turn can be written as the sum of 2 terms, each of the form $n_1 n_2$ with $n_1 \geqslant a_1$, $n_2 \geqslant a_2$. By distributivity we obtain a sum of 4 terms of the required form. ∎

Proof of Proposition 4.5. If we take a closer look at the proof of (4.5), using Theorem 2.1 and the fact that a regular polygon with n_1 vertices is an averaging set of degree p on the circle S^1 whenever $n_1 \geqslant p + 1$, we see that there exist averaging sets of degree p on S^d of size $n_1 n_2 \cdots n_d$ for any choice of integers $n_1 \geqslant p+1$ and $n_i \geqslant C_i p^{i+1}$ $(i = 2, ..., d)$, where the C_i are constants independent of p. Since a disjoint union of averaging sets is an averaging set and one can always slightly rotate averaging sets on the sphere to avoid a finite set of points, Lemma 4.6 implies the existence of an averaging set of degree p on S^d of size n for all $n \geqslant Cp \prod_{i=2}^{d} p^{i+1} = Cp^{d^2/2+3d/2-1}$, where C is independent of p. ∎

REFERENCES

1. J. ARIAS DE REYNA, A generalized mean-value theorem, *Monatsh. Math.* **106** (1988), 95–97.
2. B. BAJNOK, Construction of spherical t-designs on S^2, preprint.
3. S. N. BERNSTEIN, Sur les formules de quadrature de Cotes et Tchebycheff, *C. R. Acad. Sci. URSS* **14** (1937), 323–326. [Reprinted in "Collected Works," Vol. II, pp. 200–204, Izdat. Akad. Nauk SSSR, Moscow, 1954 (Russian).]
4. R. P. BOAS, Inequalities for the derivatives of polynomials, *Math. Mag.* **42** (1969), 165–174.
5. F. COSTABILE, Sulle formule di quadratura di Tschebyscheff, *Calcolo* **11** (1974), 191–200.
6. P. DELSARTE, J. M. GOETHALS, AND J. J. SEIDEL, Spherical codes and designs, *Geom. Dedicata* **6** (1977), 363–388.
7. W. GAUTSCHI, Advances in Chebyshev quadrature, *in* "Numerical Analysis" (G. A. Watson, Ed.), Lecture Notes in Math., Vol. 506, Springer-Verlag, Berlin/Heidelberg/New York, 1976.
8. A. MEIER AND A. SHARMA, A variation on the Tchebicheff quadrature problem, *Illinois J. Math.* **11** (1967), 535–546.
9. P. SEYMOUR AND T. ZASLAVSKY, Averaging sets: A generalization of mean values and spherical designs, *Adv. in Math.* **52** (1984), 213–240.
10. W. SIERPINSKI, "Elementary Theory of Numbers" (A. Schinzel, Ed.), Polish Sci. Publ., Warszawa, 1987.
11. A. H. STROUD, "Approximate Calculation of Multiple Integrals," Prentice–Hall, Englewood Cliffs, NJ, 1971.
12. G. SZEGŐ, "Orthogonal Polynomials," Amer. Math. Soc. Colloquium Publ., Vol. 23, 1939.