PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 41, Number 2, December 1973

SUMS OF DISTANCES BETWEEN POINTS ON A SPHERE. II

KENNETH B. STOLARSKY

ABSTRACT. Given N points on a unit sphere in Euclidean m space, $m \geqq 2$, we show that the sum of all distances which they determine plus their discrepancy is a constant. As applications we obtain (i) an upper bound for the sum of the distances which for $m \geqq 5$ is smaller than any previously known and (ii) the existence of N point distributions with small discrepancy. We make use of W. M. Schmidt's work on the discrepancy of spherical caps.

1. **Introduction.** For $m \geqq 2$ let d be a function on $U \times U$ where $U = U^m$, the surface of the unit sphere of m dimensional Euclidean space E^m. Let M_p be a sequence of N points $p_1, \cdots, p_N \in U^m$. Define

$$S(N, m, M_p) = S(d; N, m, M_p) = \sum_{i<j} d(p_i, p_j) \tag{1.1}$$

and

$$S(N, m) = S(d; N, m) = \max S(d; N, m, M_p) \tag{1.2}$$

where the maximum is taken over all sequences M_p. We wish to obtain estimates for $S(N, m)$. Our main result, Theorem 2 of §2, shows in a very exact sense that for a certain class of functions d, *including the usual Euclidean distance* $d(p, q) = |p - q|$, the quantity $S(N, m, M_p)$ is large or small depending upon whether $D(M_p)$, the discrepancy of M_p, is small or large. Since W. M. Schmidt **[9]** has obtained very good results on the discrepancy of point distributions on U, we can obtain (see Theorem 1 below) estimates on $S(N, m)$ which are far better than any hitherto known. Earlier results can be found, *inter alia*, in **[1]**, **[2]**, **[3**, p. 261, Remark 1], **[4]**, **[5]**, **[8**, pp. 36–38], and **[12]**; however, this paper can be read independently of these. Also, by applying Lemma 2.4 of **[1]** we can show (see Theorem 3 of §4) that there exist finite sequences M_p having small discrepancy. A precise definition of the discrepancy $D(M_p)$ is given after the statement of Theorem 2.

If d_0 is the great circle metric then $S(d_0; N, m) = (\pi/4)N^2$ for N even (see

Presented to the Society, December 18, 1972; received by the editors February 2, 1973.
AMS (MOS) subject classifications (1970). Primary 52A25, 52A40.

[7] and [11]; [7] also discusses the case N odd). For $q_1, q_2 \in U$ define

$$d_1(q_1, q_2) = \sigma(U)^{-1}\int |d_0(q_1, p) - d_0(q_2, p)|\, d\sigma(p) \tag{1.3}$$
$$\leqq d_0(q_1, q_2) \leqq (\pi/2)\,|q_1 - q_2|$$

where $d\sigma(p)$ is the element of surface area on U, and the total surface area of U is denoted by $\sigma(U)$. Clearly d_1 is a metric. Now (and henceforth) we let $c(m)$, $c_1(m)$, $c_i(m, \varepsilon)$, etc., denote positive constants depending on the indicated parameters. Here $c(m)$ and $c_1(m)$ equal the coefficient of N^2 in (2.4), for the appropriate ρ.

THEOREM 1. *For $\varepsilon > 0$ and $m \geqq 3$ we have*

$$c_1(m)N^2 - c_2(m)N^{1-1/(m-1)} < S(d_1; N, m) \tag{1.4}$$
$$< c_1(m)N^2 - c_3(m, \varepsilon)N^{\theta_1(m)-\varepsilon}$$

where $\theta_1(m) = (m^2 - 4m + 2)(m-1)^{-2}$. Similarly, when d is the Euclidean metric and $m \geqq 3$ we have

$$c(m)N^2 - c_4(m)N^{1-1/(m-1)} < S(d; N, m) \tag{1.5}$$
$$< c(m)N^2 - c_5(m, \varepsilon)N^{\theta(m)-\varepsilon}$$

where $\theta(m) = (m^2 - 5m + 2)(m-1)^{-2}$.

The older results, when applied to d, yield at best

$$S(d; N, m) < c(m)N^2 - c'(m) \tag{1.6}$$

for some constant $c'(m) > 0$. In fact, aside from [1], they yield results weaker than (1.6). Thus for $m \geqq 5$ the result of (1.5) is stronger than any previously known. It suggests the following question: is there a positive constant $h(m)$ such that

$$S(N, m) = c(m)N^2 - h(m)N^{1-1/(m-1)} + o(N^{1-1/(m-1)})?$$

2. **The main result.** We let $p_0 \in U$ denote the vector $(1, 0, \cdots, 0)$ and we denote the inner product of vectors $p, q \in U$ by $p \cdot q$. Note that $p \cdot \tau q = \tau^{-1}p \cdot q$ for any orthogonal transformation $\tau \in SO = SO(m)$, the special orthogonal group acting on U^m. We let $\int \cdots d\tau$ denote a normalized Haar integral over this group. For integrals on the real line we let $\int_{a,b}$ denote $\int_a^b$ unless $b < a$, in which case it shall denote $\int_b^a$.

DEFINITION 2.1. For $p_1, p_2 \in U$ and a function $g = g(x)$ integrable in $[0, 1]$ we define

$$\rho(p_1, p_2) = \rho(g; p_1, p_2) = \iint_{\tau p_1 \cdot p_0,\, \tau p_2 \cdot p_0} g(x)\, dx\, d\tau. \tag{2.1}$$

Clearly ρ is independent of p_0, and for $\tau_1 \in SO$ we have

$$\rho(\tau_1 p_1, \tau_1 p_2) = \rho(p_1, p_2). \tag{2.2}$$

Note that ρ is a metric if $g(x)>0$; however, our proof of Theorem 2 does not require this hypothesis. We call g the kernel of ρ. If $g(x)=1$ then $\rho(p_1, p_2)$ is a constant multiple of the Euclidean distance $|p_1-p_2|$. The metric d_1 of §1 has kernel $(1-x^2)^{-1/2}$. It is sometimes useful to relax the restriction on g to integrability on closed subintervals of the open interval (0, 1) and consider ρ to be defined whenever the inner integral is integrable over SO. In particular this permits the kernel $(1-x^2)^{-1}$.

DEFINITION 2.2. Let $\sigma(x)=\{p \in U \mid p \cdot p_0 \leqq x\}$. We also denote the surface area of this set by the *same* symbol. Thus

$$\sigma(x) = \int_{p\cdot p_0 \leqq x} d\sigma(p) \quad \text{and} \quad \sigma(1) = \sigma(U). \tag{2.3}$$

Set $\sigma^*(x)=\sigma(x)/\sigma(U)$. Next, let $f_p=f(M_p, \tau, x)=|M_p \cap \tau\sigma(x)|$ where $|T|$ denotes the cardinality of the set T. Thus f_p is the number of points of M_p which lie in a certain spherical cap congruent to $\sigma(x)$.

THEOREM 2.

$$\begin{aligned} S(\rho; N, m, M_p) + \int_{-1}^{1} g(x) \int (f(M_p, \tau, x) - N\sigma^*(x))^2 \, d\tau \, dx \\ = N^2 \cdot 2^{-1}\sigma(U)^{-2} \iint \rho(p, q) \, d\sigma(p) \, d\sigma(q). \end{aligned} \tag{2.4}$$

This is our main result. The second term on the left of (2.4) clearly measures the discrepancy of M_p; i.e. the extent to which it deviates from a uniform distribution. We denote it by $D(M_p)$, and call it the "discrepancy of M_p with respect to the weight (or kernel) $g(x)$". Theorem B of **[9]** shows that, for $\varepsilon>0$,

$$D(M_p) \gg N^{1-1/(m-1)-\varepsilon} \tag{2.5}$$

for *any* M_p when $g(x)=(1-x^2)^{-1}$; the implied constant depends on ε. Theorem 2 is perhaps best appreciated as an invariance principle: the sum of all distances determined by N points plus their discrepancy is constant.

3. **The proof.** First we prove a useful identity, various forms of which have already appeared explicitly or implicitly in the literature; see **[1]**, **[2]**, **[6**, p. 196, Theorem 3.1], **[10]**, and **[12]**.

LEMMA. *If p_i and q_j are real numbers for $1 \leqq i \leqq u$ and $1 \leqq j \leqq v$ with $p_1 \leqq \cdots \leqq p_u$ and $q_1 \leqq \cdots \leqq q_v$ then*

$$\sum_{i,j} \int_{p_i, q_j} - \sum_{i<j} \int_{p_i, p_j} - \sum_{i<j} \int_{q_i, q_j} = \int_{-\infty}^{\infty} g(x) G(G - (u - v)) \, dx \tag{3.1}$$

where $\int_{a,b} = \int_{a,b} g(x)\,dx$ *and*

$$G = \sum_{p_j \leqq x} 1 - \sum_{q_j \leqq x} 1. \tag{3.2}$$

Note. If $u=v$ and $g(x)=1$ we have

$$\int_{-\infty}^{\infty} G^2\,dx \geqq \int_{-\infty}^{\infty} |G|\,dx = \sum_{i=1}^{u} |p_i - q_i|$$

and (3.1) becomes the inequality of Lemma 2.1 in Alexander's paper **[1]**.

PROOF. Let x be a real number distinct from any p_i or q_j. Let s be the number of p_i to the left of x and t the number of q_j to the left of x. From the identity

$$s(v-t) + t(u-s) - s(u-s) - t(v-t) = (s-t)((s-t)-(u-v))$$

we see that $g(x)$ occurs on the left of (3.1) the same net number of times as it occurs on the right. This proves the Lemma.

We begin our proof of Theorem 2 by letting $u=v=N$ and letting M_p and M_q be the sequences $p_1, \cdots, p_N$ and $q_1, \cdots, q_N$ respectively. We let ϕ_1, ϕ_2, and τ be elements of SO and

$$\sum = \sum_{i,j} \rho(\phi_1 p_i, \phi_2 q_j).$$

If R denotes the right side of (2.4) then

$$2R = \iint \sum d\phi_1\,d\phi_2. \tag{3.3}$$

We now recall (2.1), apply the above lemma to $\sum$ with p_i and q_j replaced by $\tau\phi_1 p_i \cdot p_0$ and $\tau\phi_2 q_j \cdot p_0$ respectively, and then apply (2.2). This, and an integration over τ, yields

$$\begin{aligned} \sum = &\sum_{i<j} \rho(p_i, p_j) + \sum_{i<j} \rho(q_i, q_j) \\ &+ \iint_{-1}^{1} g(x)\Big(\sum_{\tau\phi_1 p_i \cdot p_0 \leqq x} 1 - \sum_{\tau\phi_2 q_i \cdot p_0 \leqq x} 1\Big)^2 dx\,d\tau. \end{aligned} \tag{3.4}$$

Upon inserting (3.4) into (3.3) we obtain

$$\begin{aligned} 2R = &\, S(N, m, M_p) + S(N, m, M_q) \\ &+ \int_{-1}^{1} \iint g(x)(f(M_p, \phi_1, x) - f(M_q, \phi_2, x))^2\,d\phi_1\,d\phi_2\,dx. \end{aligned} \tag{3.5}$$

Now clearly

$$\int (f_p - N^*)\,d\phi_1 = \int (f(M_p, \phi_1, x) - N\sigma^*(x))\,d\phi_1 = 0 \tag{3.6}$$

since $N^*=N\sigma^*(x)$ is just the expected value of $f_p=f(M_p, \phi_1, x)$. Since

$$(3.7)\quad (f_p - f_q)^2 = (f_p - N^*)^2 - 2(f_p - N^*)(f_q - N^*) + (f_q - N^*)^2$$

and the integral of the middle term on the right of (3.7) is zero, the right side of (3.5) is the sum of the discrepancies of M_p and M_q with respect to $g(x)$. The proof is completed by setting $p_i=q_i$ for $1\leqq i\leqq N$ and dividing both sides of (3.5) by 2.

Note. One could obtain a result having a more general appearance than (2.4) by using the full strength of the lemma rather than the special case $u=v$. A comparison of this result with (2.4) yields the identity

$$(3.8)\qquad \int_{-1}^{1} g(x)\sigma(x)(\sigma(U) - \sigma(x))\,dx = \frac{1}{2}\iint \rho(p, q)\,d\sigma(p)\,d\sigma(q).$$

4. **Some applications.** We first introduce some notation from [9]. Let μ be the normalized Lebesgue measure on U; thus $\mu(U)=1$. Let $C=C(r,p)$ be the spherical cap consisting of all points whose spherical distance from $p\in U$ is at most r. Note that $0\leqq r\leqq\pi$. Let $\Delta=\Delta(r,p)=N\mu(C)-\nu(C)$ where $\nu(C)$ denotes the number of $p_1,\cdots,p_N$ which lie in C. Let $E(r)=\int_U \Delta^2\,d\mu(p)$. Schmidt's Theorem B is given on p. 59 of [9]. By replacing his n and δ with $m-1$ and $\pi/2$ respectively we obtain the

THEOREM. *If $m\geqq 3$, $\varepsilon>0$, and $N>\varepsilon$ then*

$$(4.1)\qquad \int_0^{\pi/2} r^{-1}E(r)\,dr \geqq c_6(m,\varepsilon)N^{1-1/(m-1)-\varepsilon}.$$

For later use we note the trivial estimate

$$(4.2)\qquad E(r) \leqq c_7(m)N^2r^{m-1}.$$

(This is Lemma 4, p. 67 of [9]; set $s=r$ and note that Schmidt's $E(r,r)$ is our $E(r)$.) We now obtain a result which shows that (4.1) is not too far from best possible in its dependence on N. For $m\geqq 3$ we have

THEOREM 3. *There are points $p_1,\cdots,p_N$ on U such that*

$$(4.3)\qquad \int_0^{\pi/2} r^{-1}E(r)\,dr \leqq c_8(m)N^{1-1/(m-1)}.$$

To prove this we need the following special case of a result of Alexander (Lemma 2.4 of [1]).

LEMMA. *Let $\rho(p_1,p_2)$ be a nonnegative function on $U\times U$ and let μ be a Borel measure for which $\mu(U)=1$. Let $P=\{A_1,\cdots,A_N\}$ be a collection of Borel subsets of U such that $\mu(A_i\cap A_j)=\delta_{ij}N^{-1}$ where δ_{ij} is 1 if $i=j$ and 0* *otherwise. Then*

$$N^2\int_U\int_U \rho(p,q)\,d\mu(p)\,d\mu(q) - \sum_{i=1}^{N}\rho(A_i) \leqq 2S(\rho; N, m) \tag{4.4}$$

where $\rho(A_i)$ *is the "diameter" of* A_i; *i.e.* $\rho(A_i)=\sup \rho(p,q)$ *for* $p, q \in A_i$ (*Alexander has* $|p_1-p_2|$ *in place of* $\rho(p_1,p_2)$ *but his proof required only the nonnegativity of this function*).

We now choose M_p so that $S(\rho; N, m, M_p)=S(\rho; N, m)$. It follows from (2.4) and (4.4) that

$$\int_{-1}^{1} g(x)\int (f(M_p,\tau,x)-N\sigma^*(x))^2\,d\tau\,dx \leqq c_9(m)N\inf_P \max_i \rho(A_i), \tag{4.5}$$

a result of interest in itself.

To prove Theorem 3 let $g(x)=(1-x^2)^{-1}$ and set $p_1\cdot p=\cos\theta_1$ and $p_2\cdot p=\cos\theta_2$ where θ_1 and θ_2 are the great circle distances between p_1, p and p_2, p respectively. Then

$$\begin{aligned}\rho = \rho(p_1,p_2) &= c_9(m)\iint_{p_1\cdot p,\, p_2\cdot p} g(x)\,dx\,d\sigma(p)\\ &= c_{10}(m)\int |\log\tan\tfrac{1}{2}\theta_1-\log\tan\tfrac{1}{2}\theta_2|\,d\sigma(p).\end{aligned}$$

It is easy to see that this integral is finite since $\int |\log\tan\frac{1}{2}\theta_1|\,d\sigma(p)$ is finite. Now

$$\begin{aligned}|p_1-p_2|^{-1}\rho(p_1,p_2) &= c_{10}(m)\int\{|\theta_1-\theta_2|^{-1}(\log\tan\tfrac{1}{2}\theta_1-\log\tan\tfrac{1}{2}\theta_2)\}\\ &\qquad\cdot\{|\theta_1-\theta_2|\cdot|p_1-p_2|^{-1}\}\,d\sigma(p)\\ &\leqq c_{11}(m)\int|\theta_1-\theta_2|^{-1}(\log\tan\tfrac{1}{2}\theta_1-\log\tan\tfrac{1}{2}\theta_2)\,d\sigma(p)\\ &\equiv\int_U.\end{aligned}$$

We will show that $\int_U$ remains bounded as $p_2\to p_1$. Let the great circle distance between p_1 and p_2 be $\gamma\pi/10$ where γ is a small positive parameter. Let $B_1=B_1(\gamma)$ be the set of points $p\in U$ such that $\gamma\pi\leqq\theta_i\leqq(1-\gamma)\pi$ for $i=1,2$. Then B_2, the complement of B_1, consists of two components, say B_3 and B_4. The estimates we need for the integrals over B_3 and B_4 will be essentially identical, so we give details only for B_4 which we take to be the component for which $\theta_1,\ \theta_2\leqq 2\gamma\pi$. Write $B_4=B_5\cup B_6$ where B_5, B_6 are disjoint and B_5 consists of those points of B_4 for which $\theta_1\leqq 2\theta_2$. Then

$$\int_U=\int_{B_1}+\int_{B_3}+\int_{B_5}+\int_{B_6}\ll\int_{B_1}+\int_{B_5}+\int_{B_6}.$$

The following two estimates make use of the fact that $d\sigma(p) \ll \theta_i^{m-2}\, d\theta_i$, and require $m \geqq 3$:

$$\int_{B_5} \ll \int_{B_5} |\sin(\min(\theta_1, \theta_2))|^{-1}\, d\sigma(p) \ll \int_{B_5} \theta_1^{-1}\, d\sigma(p) \ll \int_0^{\gamma\pi} d\theta = \gamma\pi$$

and

$$\int_{B_6} \ll \int_{B_6} |\theta_2^{-1} \log \theta_2|\, d\sigma(p) \ll \int_0^{\gamma\pi} |\log \theta|\, d\theta \ll \gamma\, |\log \gamma|.$$

Also,

$$\int_{B_1} \ll \int_{B_1} (|\sin \theta_1|^{-1} + |\sin \theta_2|^{-1})\, d\sigma(p) \ll 2 \int_U |\sin \theta_1|^{-1}\, d\sigma(p) < \infty.$$

Thus $\rho(p_1, p_2) \ll |p_1 - p_2|$. Now extend the integral on the left of (4.5) only to 0, and make the change of variable $x = -\cos r$. This yields

$$\int_0^{\pi/2} r^{-1} E(r)\, dr \leqq c_{12}(m) N \inf_P \max_i \rho(A_i). \tag{4.6}$$

Now clearly one can choose the A_i so that their Euclidean diameters are $\gg\ll N^{-1/(m-1)}$ for $1 \leqq i \leqq N$. But $\rho(p_1, p_2) \ll |p_1 - p_2|$ so the result follows.

An immediate consequence of Theorem 1 and the above is that if ρ has kernel $(1-x^2)^{-1}$ then

$$\begin{aligned} c_{13}(m) N^2 - c_{14}(m) N^{1-1/(m-1)} &< S(\rho; N, m) \\ &< c_{13}(m) N^2 - c_{16}(m, \varepsilon) N^{1-1/(m-1)-\varepsilon}. \end{aligned}$$

We now prove Theorem 1. The left-hand inequalities are deduced from (4.4) as in the proof of Theorem 3. For the right-hand inequality of (1.5) we set $g(x) \equiv 1$ and note that

$$\begin{aligned} D(M_p) &\geqq c_{17}(m) \int_{-1}^{0} E(\cos^{-1} |x|)\, dx = c_{17}(m) \int_0^{\pi/2} \sin r E(r)\, dr \\ &\geqq c_{18}(m) \int_0^{\pi/2} r E(r)\, dr \geqq c_{18}(m) N^{-\alpha} \int_{N^{-\alpha}}^{\pi/2} E(r)\, dr \end{aligned} \tag{4.7}$$

for any $\alpha > 0$. From (4.1) and (4.2) we obtain

$$c_7(m) N^2 \int_0^{N^{-\alpha}} r^{m-2}\, dr + N^{\alpha} \int_{N^{-\alpha}}^{\pi/2} E(r)\, dr \geqq c_6(m, \varepsilon) N^{1-1/(m-1)-\varepsilon}. \tag{4.8}$$

Set $\alpha = 1/(m-1) + 1/(m-1)^2 + 2\varepsilon/(m-1)$. Then the term on the extreme left of (4.8) has lower order of magnitude than the term on the extreme right, so it follows that

$$D(M_p) \geqq c_{19}(m, \varepsilon) N^{1-(3+4\varepsilon)/(m-1)-2/(m-1)^2-\varepsilon}. \tag{4.9}$$

To prove (1.4) we note that in this case

$$D(M_p) \geqq c_{20}(m) \int_{N^{-\alpha}}^{\pi/2} E(r)\, dr, \tag{4.10}$$

so

$$D(M_p) \geqq c_{21}(m, \varepsilon) N^{1-(2+2\varepsilon)/(m-1)-1/(m-1)^2-\varepsilon}. \tag{4.11}$$

This completes the proof of Theorem 1.

I would like to thank J. Ralph Alexander for helpful conversations concerning this work, and also for making available to me a preprint of [1].

References

1. J. Ralph Alexander, *On the set of distances determined by n points on the* 2-*sphere*, Acta Math. Acad. Sci. Hungar. **23** (1972), 443–448.

2. J. Ralph Alexander and K. B. Stolarsky, *Extremal problems of distance geometry related to energy integrals*, Trans. Amer. Math. Soc. (to appear).

3. G. Björck, *Distributions of positive mass which maximize a certain generalized energy integral*, Ark. Mat. **3** (1956), 255–269. MR **17**, 1198.

4. L. Fejes Toth, *On the sum of distances determined by a pointset*, Acta Math. Acad. Sci. Hungar. **7** (1956), 397–401. MR **21** #5937.

5. E. Hille, *Some geometric extremal problems*, J. Austral. Math. Soc. **6** (1966), 122–128. MR **33** #6507.

6. J. B. Kelly, *Metric inequalities and symmetric differences*, Inequalities, II (Proc. Second Sympos., U.S. Air Force Acad., Colo., 1967), Academic Press, New York, 1970, pp. 193–212. MR **41** #9192.

7. F. Nielson, *Om summen af afstandene mellem n punkter pa° en Kugleflade*, Nordisk Mat. Tidskr. **13** (1965), 45–50.

8. G. Pólya and G. Szegö, *Über den transfiniten Durchmesser* (*Kapazitätskonstante*) *von ebenen und räumlichen Punktmengen*, J. Reine Angew. Math. **165** (1931), 4–49.

9. W. M. Schmidt, *Irregularities of distribution*. IV, Invent. Math. **7** (1969), 55–82. MR **39** #6838.

10. I. J. Schoenberg, *On certain metric spaces arising from Euclidean spaces by change of metric and their embedding in Hilbert space*, Ann. of Math. **38** (1937), 787–793.

11. G. Sperling, *Lösung einer elementargeometrischen Frage von Fejes Tóth*, Arch. Math. **11** (1960), 69–71. MR **22** #2935.

12. K. B. Stolarsky, *Sums of distances between points on a sphere*, Proc. Amer. Math. Soc. **35** (1972), 547–549.

Department of Mathematics, University of Illinois, Urbana, Illinois 61801