COMBINATORICA
Akadémiai Kiadó – Springer-Verlag

COMBINATORICA 10 (2) (1990) 115–136

GEOMETRIC METHODS IN THE STUDY OF IRREGULARITIES OF DISTRIBUTION

R. ALEXANDER

Received November 1, 1988

Let ν be a signed measure on $\mathbf{E}^d$ with $\nu\mathbf{E}^d = 0$ and $|\nu|\mathbf{E}^d < \infty$. Define $D_s(\nu)$ as $\sup|\nu H|$ where H is an open halfspace. Using integral and metric geometric techniques results are proved which imply theorems such as the following. **Theorem A.** Let ν be supported by a finite pointset p_i. Then $D_s(\nu) > c_d(\delta_1/\delta_2)^{1/2}\{\sum_i(\nu p_i)^2\}^{1/2}$, where δ_1 is the minimum distance between two distinct p_i, and δ_2 is the maximum distance. The number c_d is an absolute dimensional constant. (The number .05 can be chosen for c_2 in Theorem A.) **Theorem B.** Let $\mathbf{D}$ be a disk of unit area in the plane $\mathbf{E}^2$, and $p_1, p_2, \ldots, p_n$ be a set of points lying in $\mathbf{D}$. If $\mathbf{m}$ if the usual area measure restricted to $\mathbf{D}$, while $\gamma_n p_i = 1/n$ defines an atomic measure γ_n, then independently of γ_n, $nD_s(\mathbf{m}-\gamma_n) \geq .0335n^{1/4}$. Theorem B gives an improved solution to the Roth "disk segment problem" as described by Beck and Chen. Recent work by Beck shows that $nD_s(\mathbf{m}-\gamma_n) \geq cn^{1/4}(\log n)^{-7/2}$.

1. Introduction

Although difficult to define precisely, the term *discrepancy* (or *irregularity of distribution*) generally refers to an estimate of how well or how poorly a class of signed measures approximates the zero measure. In nearly all instances the positive parts of these measures are some fixed uniform measure while the negative parts are varying atomic measures. This is easily explained by the close connection with the study of uniform distribution of sequences, a topic which was greatly influenced by the fundamental work of H. Weyl [18].

For a classical example, let $\{r_1, r_2, \ldots\}$ denote a sequence of numbers in the unit interval. Define the purely atomic measure γ_n of total mass 1 by placing an atom of weight $1/n$ at each point of the set $\{r_1, r_2, \ldots, r_n\}$. Letting $\mathbf{m}$ denote Lebesgue measure on the unit interval, the problem is to estimate how small $nD(n)$ can remain if

$$D(n) = \sup_I |(\mathbf{m} - \gamma_n)I|; \qquad I = [a,b], \quad I \subset [0,1].$$

T. van Aardenne–Ehrenfest showed that for any sequence $\{r_i\}$, $\limsup nD(n) = \infty$, and K. F. Roth in a famous article [9] improved this by showing that $nD(N) > c\sqrt{\log n}$ for infinitely many n. Finally, W. M. Schmidt [13] proved that $nD(n) > c\log n$ for infinitely many n. The well-known example of the Van der Corput sequence satisfies $nD(n) < c\log n$. Thus Schmidt achieved the optimal order of magnitude.

AMS subject classification (1980): 52 A 22, 10 K 30.

A crucial aspect of this problem is that the measures γ_n are defined by a fixed sequence. It is clear that for a given n, the points $\{r_1, r_2, \dots, r_n\}$ could be arranged so that $nD(n) = 1$. However, a key point of Roth's method is that the problem of sequential discrepancy is shown to be equivalent to a certain discrepancy problem involving the free placement of n points in a unit square. This important advance opened the door for consideration of a much wider class of problems in discrepancy theory.

The book of Kuipers and Neiderreiter [8], which is devoted to the study of uniform distribution, contains discussions of the classical problems of discrepancy, and details some of the methods of Roth and Schmidt. The very recent book by Beck and Chen [7] contains a much more extensive treatment of these topics and, in addition, a great number of fine new ideas and theorems.

In this paper we will investigate what is termed the *separation discrepancy* D_s. Let ν be any signed Borel measure on E^t such that $\nu E^t = 0$. Define $D_s(\nu)$ by

$$D_s(\nu) = \sup |\nu H|, \quad H \text{ an open halfspace.}$$

A number of problems appearing in the literature can be expressed in terms of this concept.

The method of investigation used in the present work involves a combination of integralgeometric and metricgeometric techniques. The results obtained stand up well to comparison with those obtained by the various analytic methods. The fact that the central ideas are principally geometric in nature allows a clearer perception of certain aspects of discrepancy that might remain obscured by some other methods. The following theorems give an idea of the type of lower bound inequalities that can be achieved for D_s.

Theorem A. *Let ν be a totally atomic measure on E^t concentrated on the points $p_1, p_2, \dots, p_n$ such that $\sum_i \nu p_i = 0$. Then the separation discrepancy satisfies*

$$D_s(\nu) \geq c_t(\delta_1/\delta_2)^{1/2}\Big\{\sum_i (\nu p_i)^2\Big\}^{1/2}, \tag{1}$$

where δ_1 is the minimum distance between two distinct p_i, and δ_2 is the maximum distance, or the diameter of the point set. The number $c_t > 0$ is an absolute constant.

One interesting situation where Theorem A applies is if the measure ν is concentrated at integral lattice points. For example, suppose that ν assigns measure ± 1 to the $N = n^t$ lattice points contained in a lattice t-cube of edge length $n-1$. A straightforward calculation shows that Theorem A implies that $D_s(\nu) \geq c_t t^{-1/4} N^{1/2(1-1/t)}$.

While the study of discrepancy for discrete measures is a natural setting for the present method, it can be extended to treat other measures although a number of problems remain. We state one example.

Theorem B. *Let* $\mathbf{D}$ *be a disk of unit area in the plane E^2, and $p_1, p_2, \dots, p_n$ be a set of points lying in* $\mathbf{D}$. *If* $\mathbf{m}$ *is the usual area measure restricted to* $\mathbf{D}$ *while $\gamma_n p_i = 1/n$ defines an atomic measure γ_n, then independently of γ_n,*

$$nD_s(m - \gamma_n) \geq .0335\, n^{1/4}. \tag{2}$$

Theorem B gives an improved solution to the Roth "disk segment problem" as described by Beck and Chen [7,§7.0]. Roth conjectured that for arbitrary γ_n, $nD_s(\mathbf{m} - \gamma_n)$ is unbounded with n, and Beck [7,§7.3] confirmed the conjecture by showing that for some constant c, $cn^{1/4}(\log n)^{-7/2}$ is a lower bound. This problem also appears in W. M. Schmidt's paper [12].

The present method allows example constants c_t to be calculated without outrageous effort if need be. For example, in Theorem A the constant c_2 may be chosen as .05. However, it would be very surprising indeed if the $n^{1/4}$ in (2) indicated the exact order of magnitude of D_s for the disk segment problem. The method does yield the exact order of magnitude for certain L^2 averages closely related to D_s. There are further comments in Section 11.

The author's longstanding interest in metric geometry began with the joint paper [5] with Kenneth B. Stolarsky. Shortly thereafter appeared articles [1], [15], and [2] which all greatly influenced the present work. In [1] the author introduced a simple probabilistic method for obtaining good upper bounds on metric inequalities (see Section 11), and in [15] Stolarsky made his wellknown discovery of the close relationship between discrepancy on euclidean spheres and metric geometry (see Section 10). In [2] it was shown that distance sums possess an integral geometric formulation. These papers opened the door to the possibility that geometric ideas might contribute to the study of discrepancy phenomena.

Finally, let us state what is the major difference between the results achieved by our new method and those preceding. The precise geometry of the measure ν plays practically no role. The dimension of the support of ν is the single most important parameter. The estimates obtained in Theorem B hold for a square or more generally for a bounded measurable set on a convex 2-surface of bounded curvature. See Theorem 20 and Theorem 25.

2. The Crofton arclength formula and the functionals I and J

It is well understood that there is a motion invariant Borel measure μ on the $(t-1)$-planesets of E^t which, when suitably normalized has the property that for any points p, q in E^t,

$$|p - q| = 1/2\mu\{h : h \text{ cuts seg } pq\} \tag{3a}$$

More generally, if K is a planar convex body, then

$$\text{Perimeter } K = \mu\{h : h \text{ cuts } K\} \tag{3b}$$

Except for the plane ($t = 2$) this normalization differs from the usual normalization of the measure μ. This is discussed in Section 9. The virtue of the present normalization for the work of this article is that the measure μ on the $(t-1)$-planesets of E^t is induced in the most natural manner from the measure μ on the t-planesets of E^{t+1}. Thus one need not specify the "dimension" of μ. The constructions described in this article sometimes require that t be increased. Formula (3) and derivative formulas remain unchanged by such processes.

Letting ν be a finite signed Borel measure on E^t, the functional I is defined by

$$I(\nu) = \iint |p - q| \, d\nu(p) d\nu(q). \tag{4a}$$

If ν and $\bar{\nu}$ are two such measures, the associated bilinear functional J is defined by

$$J(\nu,\bar{\nu}) = \iint |p-q|\, d\nu(p)d\bar{\nu}(q). \tag{4b}$$

As was shown in [2] (with notational changes), the Crofton formula (3a) leads to the following representation for I as an integral with respect to the measure μ,

$$I(\nu) = \int A(h)B(h)d\mu(h). \tag{5}$$

Here $A(h) = \nu(h^+)$ and $B(h) = \nu(h^-)$, where h^+ and h^- represent the open halfspaces determined by h. If h varies over the space of oriented planes, the halfspaces h^+ and h^- are well-defined, however, the product is well-defined for unoriented planes. As is easily shown, for almost all planes h, $A(h)+B(h) = \nu(E^t)$.

The key observation in demonstrating formula (5) is that

$$\iint |p-q|d\nu(p)d\nu(q) = 1/2 \iiint \chi(p,q,h)d\mu(h)d\nu(p)d\nu(q).$$

Here $\chi(p,q,h) = 1$ if the hyperplane h intersects the open segment $\overline{pq}$ in precisely one point and $\chi(p,q,h) = 0$, otherwise. At this stage one takes the μ-integral to the outside and then performs the $p-q$ integration. Note that for a fixed h, $\chi(p,q,h) = 1$ precisely if (p,q) lies in $\nu^+(h)\times\nu^-(h)$ or (p,q) lies in $\nu^-(h)\times\nu^+(h)$. See the article [2] for more details.

If we restrict our attention to Borel measures ν with the property that $\nu(E^t) = 0$, then formula (5) takes the shape

$$I(\nu) = -\int A(h)d\mu(h). \tag{5a}$$

If $\nu(E^t) = \bar{\nu}(E^t) = 0$, then similar ideas at once give the corresponding formula for J,

$$J(\nu,\bar{\nu}) = -\int A^2(h)\bar{A}(h)d\mu(h). \tag{5b}$$

Equation (5a) quantifies the negative semidefiniteness of the functional I. Actually, I is negative definite, i.e., if $I(\nu) = 0$, then $\nu = 0$. This fact is closely related to the uniqueness problem for the Radon transform.

If τ is a motion of E^t and ν is a measure on E^t, then the measure $\tau\nu$ is defined by $\tau\nu C = \nu\tau^{-1}C$ for Borel sets C. There is the relation

$$I(\nu) = I(\tau\nu). \tag{7}$$

Equation (7) depends not only on the definition of $\tau\nu$, but also on the motion invariance of the measure μ on the planesets. If μ were to define a Minkowski metric via equation (3), then (7) would hold generally only if τ were a translation.

For a real scalar c the measure $c\nu$ is of course defined by $c\nu C = c(\nu C)$. Equation (5) implies the relations

$$I(c\nu) = c^2 I(\nu) \text{ and } J(c\nu, c'\nu') = cc'J(\nu,\nu'). \tag{8}$$

the relation $I(-\nu) = I(\nu)$ is an important example of (8).

Next if $\Phi = \Phi^+ - \Phi^-$, set $|\Phi| = \Phi^+ + \Phi^-$ and $\|\Phi\| = |\Phi|E^t$. Let Ψ denote the class of signed Borel measures $\{\nu : \nu(E^t) = 0, \|\nu\| < \infty\}$, and let $\Psi[C]$ denote those members of Ψ with support in the set C.

Theorem 1. *The functional $-I$ is nonnegative convex on Ψ.*

Proof. Let $a_1, \dots, a_k$ be nonnegative numbers whose sum is 1. Consider

$$-I\Big\{\sum_i a_i\nu_i\Big\} = \int\Big\{\sum_i a_iA_i(h)\Big\}^2 d\mu(h) \le \Big(\sum_i a_i\Big)\int\sum_i a_i\{A_i(h)\}^2 d\mu(h);$$

the inequality is obtained by setting $a_i = \sqrt{a_i}\sqrt{a_i}$ and applying Cauchy–Schwarz. Thus

$$0 \le -I\Big\{\sum_i a_i\nu_i\Big\} \le -\sum_i a_iI(\nu_i). \tag{9}$$

∎

It should be noted that if all the measures ν_i are congruent (up to sign) to a fixed measure ν, then the right side of (9) collapses to $-I(\nu)$. Also, since the separation discrepancy satisfies $D_s(\nu) = \sup|A(h)|$, the close relation of this concept to the functional I is apparent.

Corollary 2. *Let $a_1, a_2, \dots, a_k$ be numbers such that $\sum_i |a_i| = 1$. Then*

$$-I\Big\{\sum_i a_i\nu_i\Big\} \le -\sum_i |a_i| I(\nu_i). \tag{10}$$

Proof. If $a_i < 0$, set $a_i\nu_i = |a_i|(-\nu_i)$ and apply (9), and since $I(-\nu_i) = I(\nu_i)$, the proof is complete. ∎

Also, the wellknown notion of the *convolution* $\Phi_1 * \Phi_2$ of two signed measures is helpful. For example, if Φ_1, Φ_2 are atomic measures on E^t with $\Phi_1(p_i) = a_i$ and $\Phi_1(q_j) = b_j$, then the atomic measure $\Phi_1 * \Phi_2$ with support on the set $\{p_i + q_j\}$ is defined by $\Phi_1 * \Phi_2(p_i + q_j) = a_ib_j$. Clearly, $\Phi_1 * \Phi_2(E^t) = \Phi_1(E^t)\Phi_2(E^t)$ so that the class of measures Ψ is closed under convolution. In fact it can be quickly verified that the signed Borel measures form an algebra over the real (or complex) numbers with respect to the operations $+$, $*$ and that Ψ is an ideal in this algebra.

Corollary 3. *Let ν be in Ψ, and let Φ be a signed Borel measure. Then*

$$-I(\Phi * \nu) \le -\|\Phi\|^2 I(\nu). \tag{11}$$

Proof. Since $c(\Phi*\nu) = (c\Phi*\nu)$ and $I(c\omega) = c^2I(\omega)$, it suffices to assume that $\|\Phi\| = 1$. Let $\Phi(p_i) = a_i$ where Φ is concentrated on the points $\{p_i\}$. Now $\Phi * \nu = \sum_i a_i\nu_i$ where the equation $\nu_i\chi = \nu(\chi - p_i)$ defines the measure ν_i. To complete the proof apply (10) together with the fact that $I(\nu_i) = I(\nu)$ for each i. The inequality (11) for nonatomic measures Φ follows easily, but in this article only atomic measures Φ are needed. ∎

3. The functional $\mathbf{I}^\alpha$

For a nonnegative number α and signed Borel measure ν on E^t the functional I^α is defined by

$$I^\alpha(\nu) = \iint |p - q|^\alpha d\nu(p)d\nu(q).$$

In the case where ν is area measure restricted to a planar convex body the numbers $I^k(\nu)$ $k = 0, 1, \dots$ have a long history, being studied by Crofton, Blaschke, and others. See [10] p. 46. The functional I^α for α in the range $0 < \alpha < 2$ been studied by I. J. Schoenberg as part of his very pretty theory of metric embeddings into Hilbert space.

An old result of I. J. Schoenberg [11], slightly modified, states that I^α is a negative semidefinite functional on the space of measures Ψ for $0 \le \alpha \le 2$. (Define $I^0 = \lim I^\alpha$, $\alpha > 0$.) Of course, equation (6) establishes this for $\alpha = 1$; see equation (12) below for a corresponding result for $\alpha = 2$.

The case $\alpha = 2$ is certainly the most important for the theoretical study of metric geometry in euclidean and Hilbert spaces. The paper [4] discusses a number of the geometrical aspects of I^2. As will be shown below, $I^2(\nu) = 0$ for ν in Ψ precisely when the positive and negative parts of ν share a common centroid.

If the measure Φ in $\Psi(E^t)$ is concentrated on the distinct points $\{p_i\}$ with $\Phi(p_i) = x_i$, then

$$I^\alpha(\Phi) = \sum_{ij} |p_i - p_j|^\alpha x_i x_j \text{ where } \quad \sum x_i = 0.$$

Lemma 4. *If the atomic measure Φ lies in Ψ, then the functional $-I^0(\Phi) = \sum x_i^2$.*

Proof. Note that $\lim_{\alpha\to 0} -I^\alpha(\Phi) = -\sum_{i\neq j} x_i x_j = \left[\sum x_i^2 - (\sum x_i)^2\right]$. ∎

Lemma 5. *Let Φ be an atomic measure with support in a set of diameter δ. Then*

$$|I^\alpha(\Phi)| \le \|\dot{\Phi}\|^2 \delta^\alpha.$$

Proof. Because equation (8) also holds for I^α it may be assumed that $\|\Phi\| = 1$. Certainly a supremum for $|I^\alpha(\Phi)|$ can be approached only if all x_i agree in sign, which may as well be positive so that $\sum x_i = 1$. If for each $i \neq j$ the distance $|p_i - p_j|^\alpha$ is replaced by δ^α, then $0 < |I^\alpha(\Phi)| \le \delta^\alpha \sum_{i\neq j} x_i x_j = \delta^\alpha\left[(\sum x_i)^2 - \sum x_i^2\right] \le \delta^\alpha$. This proves the lemma. ∎

4. Atomic measures with vanishing moments

Lemma 6. *Let the atomic measure Φ belong to Ψ. Then $I^2(\Phi) = 0$ if and only if $\sum x_i p_i = 0$, or equivalently, Φ^+ and Φ^- have a common centroid.*

Proof. Verify the identity

$$-I^2(\Phi) = 2\left\|\sum x_i p_i\right\|^2 = 2\left\langle \sum x_i p_j, \sum x_i p_i \right\rangle \tag{12}$$

subject to the condition $\sum x_i = 0$. ∎

We shall not investigate the general problem of when $I^\alpha(\Phi)$ can vanish, but rather concentrate on a much simpler class of measures needed for the remainder of the paper. The next lemma generalizes Lemma 6 for this class.

Lemma 7. *Let the atomic measure Φ belong to Ψ, and suppose that Φ is supported by a line (which is identified with $\mathbf{R}$). Then the metric functionals I^{2k}, $1 \le k \le n$, simultaneously vanish at Φ if and only if the first n moments of the measure $\Phi, \sum r_i^k x_i$, simultaneously vanish.*

Proof. As defined, $I^{2n}(\Phi) = \sum_{i,j} |r_i - r_j|^{2n} x_i x_j$ where $\Phi(r_i) = x_i$. Expanding each term via the binomial theorem and summing on l from 0 to $2n$ yields

$$I^{2n}(\Phi) = \sum_l (-1)^l c(2n,l)\Big(\sum r_i^l x_i\Big)\Big(\sum r_i^{2n-l} x_i\Big) \tag{13}$$

Equation (13) immediately shows that vanishing moments imply vanishing functionals I^{2k} since each term of the sum involves a moment of order not exceeding n. Note that the terms for $l = 0$ and $l = 2n$ vanish because Φ belongs to $\Psi[\mathbf{R}]$.

If the functionals $i^{2k}(\Phi), 1 \le k \le n$, vanish one may proceed by induction to show that the first n moments vanish. Lemma 6 treats the case $n = 1$. If the statement is true for $n-1$, then by assumption the first $n-1$ moments vanish. Since $I^{2n}(\Phi) = 0$, equation (13) reduces to $0 = (-1)^n(2n,n)(\sum r_i^n x_i)^2$. Thus the n-th moment vanishes. ∎

Remark. It the signed measure $\Phi \in \Psi[\mathbf{R}]$ is symmetric with respect to the origin, then all odd moments vanish. If the measure is skew symmetric instead, all even moments will vanish. We give two examples that will be used for explicit estimates later.

Example 1. Let c be positive. Define the measure Φ_1 on $\mathbf{R}$ by $\Phi_1(-1/4) = -1/4$, $\Phi_1(0) = 1/2$ and $\Phi_1(1/4) = -1/4$. Note that $\Phi_1 \in \Psi, \|\Phi_1\| = 1$ and $\sum r_i x_i = 0$.

Example 2. Let c be positive, set $K = (2/c + 2c)^{-1}$. Define the measure Φ_2 on $\mathbf{R}$ by $\Phi_2(-c) = -K/c, \Phi_2(-1/c) = Kc, \Phi_2(1/c) = -Kc, \Phi_2(c) = K/c$. Note that $\Phi_2 \in \Psi, \|\Phi_2\| = 1, \sum r_i x_i = 0$ and $\sum r_i^2 x_i = 0$.

Theorem 8. *For any given n there is an atomic measure $\Phi \in \Psi[-c,c]$, concentrated on at most $2n+2$ atoms of equal weight, such that the functionals $I^2, I^4, \ldots, I^{2n}$ all vanish at Φ.*

Proof. Let $p = (r_1, \ldots, r_{n+1})$ be a point in the positive orthant of E^{n+1} such that $r_i \ne r_j$ if $i \ne j$. For each $k, 1 \le k \le n$ choose a_k so that p lies on the surface $\sum r_i^k = a_k$. For this system of n equations in $n+1$ unknowns the $n \times n$ minors of the $n \times (n+1)$ Jacobian matrix at p are nonsingular Vandermonde matrices. The implicit function theorem implies that there is a curve through p of solutions to the system. Choose a point $\bar{p} \ne p$ in the positive orthant lying on this curve. If $\bar{p} = (\bar{r}_1, \ldots, \bar{r}_{n+1})$ then the equations

$$\sum r_i^k - \sum \bar{r}_i^k = 0, \quad k = 0, \ldots, n$$

prove the lemma. Here $\Phi(r_i) = 1$ and $\Phi(\bar{r}_i) = -1$. By choosing p and $\bar{p}$ of distance less than c from the origin of E^{n+1} we insure that $\Phi \in \Psi[-c,c]$. The factor $(2n+2)^{-1}$ will normalize the measure so that $\|\Phi\| = 1$. ∎

Theorem 2 is an existence result and gives no hint of an efficient way to proceed in obtaining Φ. A smoothing and polynomial approximation process quickly shows that no nonzero measure with compact support can have all of its moments vanish.

5. Developing the general method

Let us begin by saying that up to this point all of the important properties of the functional I have been deduced from equation (6). However, the euclidean metric has numerous important properties that may also be used to study I via its defining equation (4). Information obtained in this manner then may be employed to deduce properties of the right side of equation (6). The present investigation moves in this direction. In this section all measures mentioned will be atomic.

Consider E^t as the plane $z_{t+1} = 0$ in E^{t+1}, and let $\Phi \in \Psi[\mathbf{R}]$, where $\mathbf{R}$ is identified with the z_{t+1}-axis in E^{t+1}. We wish to carefully study the convolution measure $\Phi * \nu$ where ν in $\Psi(E^t)$ is supported by $p_1, \ldots, p_n$.

If the measure Φ is supported by the points $\{r_1, \ldots, r_m\}$, then the convolution measure $\Phi * \nu$ may be expressed as $\Phi * \nu = \sum_i (\nu p_i)\Phi_i$, $1 \leq i \leq n$, where Φ_i is supported by the points $\{(p_i, r_1), (p_i, r_2), \ldots, (p_i, r_m)\}$ and $\Phi_i(p_i, r_k) = \Phi(r_k)$. In this special situation the convolution measure $\Phi * \nu$ also has the structure of a product measure $\nu \times \Phi$. By either viewpoint $\Phi * \nu(p_i, r_k) = \Phi(r_k)\nu(p_i)$.

Define the symbol $J(p_i, p_j)$ by

$$\begin{aligned} J(p_i, p_j) &= J(\Phi_i, \Phi_j) \qquad\qquad \text{(as defined in 6b above)} \\ &= 1/2\{I(\Phi_i + \Phi_j) - I(\Phi_i) - I(\Phi_j)\}. \end{aligned}$$

This allows the following important representation of $I(\Phi * \nu)$.

Lemma 9. *Let $\sigma = (\Phi * \nu)$ be the atomic measure supported on the mn points $q_{ik} = (p_i, r_k)$ in the manner described above. Then*

$$\begin{aligned} I(\sigma) &= \sum_{i,j} J(p_i, p_j)\nu p_i \nu p_j \\ &= \sum_{i \neq j} J(p_i, p_j)\nu p_i \nu p_j + I(\Phi)\sum_i (\nu p_i)^2. \end{aligned} \tag{14}$$

Proof. By the definition of the functional I and because of the geometric structure of the measure σ one may write

$$I(\sigma) = \sum_{i,j}\sum_{k,l}(d_{ij}^2 + h_{kl}^2)^{1/2}\sigma(q_{ik})\sigma(q_{jl}) \tag{15}$$

where $d_{ij} = |p_i - p_j|$ and $h_{kl} = |r_k - r_l|$. Note that for a fixed pair i, j for which $i \neq j$, all the terms of the sum (15) fall into $J(p_i, p_j)\nu p_i \nu p_j$. If $i = j$ the terms fall into the sum $I[(\nu p_i)\Phi_i]$. Finally note that $I(\nu p_i \Phi_i) = (\nu p_i)^2 I(\Phi)$. These observations verify the lemma. ∎

The previous proof of Lemma 9 has the advantage of showing the role of the pythagorean theorem in calculating $I(\sigma)$, and therefore gives insight into the crucial Lemma 10 below. However we include a direct formalistic proof based on formulas

(25) and (26) below. If $a_i = \nu(p_i)$,

$$I(\sigma) = J(\sigma,\sigma) = J(\Sigma a_i\Phi_i, \Sigma a_i\Phi_i) = \sum_{ij} J(a_i\Phi_i, a_j\Phi_j)$$
$$= \sum_{i\neq j} a_i a_j J(\Phi_i,\Phi_j) + \sum_i a_i^2 J(\Phi_i,\Phi_i)$$
$$= \sum_{i\neq j} J(p_i,p_j)\nu p_i \nu p_j + I(\Phi)\sum_i \nu(p_i)^2.$$

The present method will show that by a suitable choice of Φ the sum $\sum_{i\neq j} J(p_i,p_j)$ will be of smaller magnitude than $\sum_i I(\Phi_i)$. From here inequality (11) gives a lower bound on $-I(\nu)$. This information may then be applied to the separation discrepancy $D_s(\nu)$ via equation (6).

Since the function $J(p,q)$ depends on two variables, Φ and the distance d between the two points p,q, the notation $J(\Phi,d)$ will be used whenever convenient.

Lemma 10. *Let $\Phi \in \Psi[-1/4,1/4]$ and $d \geq 2$, then*

$$J(\Phi,d) = \sum_{k\geq 1} c_k I^{2k}(\Phi) d^{-2k+1} \tag{16}$$

where the numbers c_k are the coefficients of the binomial series expansion for $(1+x^2)^{1/2} = \sum_{k\geq 0} c_k x^{2k}$.

Proof. Write $(d^2+h^2)^{1/2} = d[1+(h/d)^2]^{1/2} = d\{\sum_{k\geq 0} c_k(h/d)^{2k}\}$. In this shape sum over the m^2 terms in equation (15) that are associated with $J(p_i,p_j)$ for some fixed i,j where $d = |p_i - p_j|$. Consider the coefficient of d^{-2k+1} and note that when $k=0$ this coefficient is $(\Phi\mathsf{R})^2$ which is 0 since $\Phi \in \Psi(\mathsf{R})$. For positive k the coefficient of d^{-2k+1} will be $c_k I^{2k}(\Phi)$ because if $h_{ij} = |r_i - r_j|$, then the term $c_k h_{ij}^{2k}\Phi(r_i)\Phi(r_j)$ will appear exactly two times. ∎

We now proceed to a somewhat detailed study of the discrepancy D_s of signed atomic measures in E^2. The measure Φ_1 in Example 1 above plays a central role, and because of its simple structure more precise estimates can be made than in higher dimensions.

Lemma 11. *Let $d \geq 2$ and let Φ_1 be the measure on R given by Example 1 above. Then*

$$|J(\Phi_1,d)| < (3/4096)d^{-3}. \tag{17}$$

Proof. We begin by noting that the binomial series coefficient c_k satisfies $|c_k| \leq 1$ for all k and that c_k decreases in magnitude while alternating in sign. Now $I^2(\Phi_1) = 0$, and for $k > 1$, $I^{2k}(\Phi_1)$ is easily seen to be a monotone decreasing positive sequence since $I^{2k}(\Phi_1) = 4^{-2(k+1)}[2^{2k+1} - 8]$. Thus it follows at once from the most elementary properties of alternating series applied to equation (16) that $|J(\Phi_1,d)| < |c_2 I^4(\Phi_1)d^{-3}|$. The fact that $c_2 = -1/8$ leads at once to inequality (17). ∎

Lemma 12. *Consider the quadratic form* $\sum_{ij} a_{ij}x_ix_j$ *where* $a_{ij} \geq 0, a_{ij} = a_{ji}$. *Suppose that for each* i, $\sum_j a_{ij} \leq K$. *Then* $|\sum_{ij} a_{ij}x_ix_j| \leq K\sum_i x_i^2$.

Proof. Write $|\sum_{ij} a_{ij}x_ix_j| = |\sum_{ij}(\sqrt{a_{ij}}x_i)(\sqrt{a_{ij}}x_j)|$, which is less than $\{\sum_{ij}(a_{ij}x_i^2)\}^{1/2}\{\sum_{ij}(a_{ij}x_j^2)\}^{1/2} = \sum_{ij}(a_{ij}x_i^2)$ by the Cauchy–Schwarz inequality. Writing the last sum as $\sum_i x_i^2 \sum_j a_{ij}$ leads at once to the desired conclusion.

Lemma 13. *Let* ν *be an atomic measure concentrated on the points* p_i *having a minimum separation* $\delta_1 \geq 2$, *and let* Φ_1 *be as in Example* 1. *Then*

$$\Big|\sum_{i\neq j} J(p_i,p_j)\nu(p_i)\nu(p_j)\Big| < 1/60\Big\{\sum_i(\nu p_i)^2\Big\}. \tag{18}$$

Proof. Since $|\sum_{i\neq j} J(p_i,p_j)\nu(p_i)\nu(p_j)| \leq \sum_{i\neq j}|J(p_i,p_j)\nu(p_i)\nu(p_j)|$, in view of lemma 12 with $|J(p_i,p_j)|$ as a_{ij}, it will suffice to show that for any i,

$$\sum_j \{|J(p_i,p_j)|, i\neq j\} < 1/60. \tag{19}$$

Take a fixed member of the p_i, denoted by p_0, and let $d_1 \leq d_2 \leq \ldots$ be the distances to the other points indexed in order of increasing magnitude. Then

$$d_k > k^{1/2}/2. \tag{20}$$

The inequality (20) follows from the observation that the $k+1$ open disks of radius 1 centered at $p_0, p_1, \ldots, p_k$ are pairwise disjoint and lie within a disk of a radius $d_k + 1$; thus $\pi(n+1) < \pi(d_k+1)^2$. This inequality leads at once to (20). From (20) and (17) it follows that

$$|J(p_0,p_k)| < (24/4096)k^{-3/2}$$

and that

$$\sum_k |J(p_0,p_k)| < (24/4096)\sum_k k^{-3/2}.$$

Since $\sum_k k^{-3/2} < 2.671$, the inequality (19) follows at once, completing the proof of the lemma. ∎

Theorem 14. *Let* ν *be an atomic measure of total mass* 0 *concentrated on the points* p_i *in the plane* E^2. *Suppose that the minimum distance between any two* p_i *is at least* δ_1. *Then*

$$-I(\nu) > .02\delta_1\Big\{\sum_i(\nu p_i)^2\Big\}. \tag{21}$$

Proof. The form of the functional I allows the assumption that $\delta_1 = 2$ with no loss of generality since I is linear with respect to ratios of dilation.

Now $-I(\Phi_1 * \nu) \leq -I(\nu)$ by (11). However, by (14)

$$I(\Phi_1 * \nu) = \sum_{i\neq j} J(p_i,p_j)\nu p_i \nu p_j + I(\Phi_1)\sum_i(\nu p_i)^2.$$

Since $-I(\Phi_1) = 1/16$ and $|\sum_{i\neq j} J(p_i,p_j)\nu p_i\nu p_j| < (1/60)\sum_i(\nu p_i)^2$, by (18), it follows that $-I(\Phi_1 * \nu) > (.045)\sum_i(\nu p_i)^2$. This establishes inequality (21). ∎

Theorem 15. (Theorem A for t=2.) *Let ν be an atomic measure of total mass 0 concentrated on the points p_i in the plane E^2. Then*

$$D_s(\nu) > .05\{\delta_1\delta_2^{-1}\sum_i(\nu p_i)^2\}^{1/2} \tag{22}$$

where δ_1 and δ_2 are as previously defined.

Proof. Since for all lines h there is the inequality $A(h) \le D_s(\nu)$, it follows from Theorem 14 and equation (6) that

$$\{D_s(\nu)\}^2\mu K > .02\delta_1\Big\{\sum_i(\nu p_i)^2\Big\}, \tag{23}$$

where K is the support lineset for the function A. However, as is well understood,

$$\mu(K) \le \text{Perimeter Conv}(p_i) \le 2\pi\delta_2. \tag{24}$$

Inequality (22) follows at once from inequalities (23), (24), and the fact that $(.02/2\pi)^{1/2} > .05$. ∎

6. Further development of the method for atomic measures

The theorem of Fubini allows one to write

$$I(\alpha\nu_1+\beta\nu_2) = \alpha^2I(\nu_1)+2\alpha\beta J(\nu_1,\nu_2)+\beta^2I(\nu_2), \text{ and} \tag{25}$$

$$I(\Phi*(\nu_1+\nu_2)) = I(\Phi*\nu_1)+2J(\Phi*\nu_1,\Phi*\nu_2)+I(\Phi*\nu_2). \tag{26}$$

Before more precise estimates of discrepancy in E^2 can be made, it is necessary to take a closer look at the key function $J(\Phi_1, d)$ where J and Φ_1 are as defined above. A straightforward calculation gives

$$16\, J(\Phi_1,d) = 6d + 2(d^2+1/4)^{1/2} - 8(d^2+1/16)^{1/2}.$$

Elementary calculus shows that $\lim_{d\to\infty} J = \lim_{d\to\infty} J' = 0$. Also $J' > 0$ and $J'' < 0$ if $d > 0$. These facts at once imply the following lemma.

Lemma 16. *The function $-J(\Phi_1,d)$ is a monotone decreasing positive function of d for $d \ge 0$. Also $-J(\Phi_1,0) = -I(\Phi_1) = 1/16$).*

To illustrate how the nonvanishing property of J applies, we prove a result which in some situations will allow a weaker hypothesis on the minimal distances between the atoms of ν. Later the full strength of Lemma 16 will be needed for applications to measures with continuous parts.

If ν is an atomic measure on $\Psi(E^t)$ supported by the points p_i, let δ_1^+ denote the minimum distance between two points both having positive measure, δ_1^- denote the minimum distance between two points both having negative measure, and $\delta^\#$ denote the minimum distance between two points of opposite measure. Set

$$\bar{\delta} = \min\{\max\{\delta_1^+,\delta_1^-\},\delta^\#\}.$$

Theorem 17. *Let ν be an atomic measure of total mass 0 concentrated on the points p_i, $i = 1, \dots, 2n$ in the plane E^2, and suppose that $|\nu p_i| = c$ for each i. Then*

$$D_s(\nu) > .05(\bar{\delta}/\delta_2)^{1/2} c\sqrt{2n} \tag{27}$$

where $\bar{\delta}$ and δ_2 are as previously defined.

Proof. We will show that

$$- I(\nu) > .045\, \bar{\delta}\, nc^2, \tag{28}$$

and inequality (27) follows from the same argument used to derive inequality (22) from inequality (21). If $\nu = \nu^+ - \nu^-$, let $\nu_1 = \Phi_1 * \nu^+$ and $\nu_2 = \Phi_1 * \nu^-$. Thus $\Phi_1 * \nu = \nu_1 - \nu_2$, and applying (26) gives

$$-I(\Phi_1 * \nu) = -I(\nu_1) + 2J(\nu_1, \nu_2) - I(\nu_2).$$

Applying (14) to $I(\nu_1)$ gives

$$I(\nu_1) = \sum_{i \neq j} J(p_i, p_j)(\nu^+ p_i)(\nu^+ p_j) + I(\Phi_1) \sum_i (\nu^+ p_i)^2.$$

A similar relation holds for $I(\nu_2)$ by changing ν^+ to ν^-. The key point here is that $J(p_i, p_j)$ and $I(\Phi_1)$ are both negative numbers. This follows from Lemma 16. Thus it may be concluded that

$$\begin{aligned} -I(\nu_1) - I(\nu_2) &> -I(\Phi_1) \sum_i (\nu^+ p_i)^2 - I(\Phi_1) \sum_1 (\nu^- p_i)^2 \\ &= I(\Phi_1) \sum_i (\nu p_i)^2 \\ &= 1/16 \sum_i (\nu p_i)^2 \\ &= nc^2/8. \end{aligned} \tag{29}$$

The next step is to obtain a bound for the number $2J(\nu_1, \nu_2)$. Without loss of generality let it be assumed that $\max\{\delta_1^+, \delta_1^-\} = \delta_1^+$. As before the linearity of the functional I with respect to dilations allows the assumption that $\bar{\delta} = 2$. With these assumptions write

$$J(\nu_1, \nu_2) = \sum_{i,j} J(p_i, p_j)(\nu^+ p_i)(\nu^- p_j),$$

and following the exact same pattern used in the proof of Lemma 13, $|\sum_i J(p_i, p_j)| < 1/60$, where p_j is a fixed atom of ν^- and p_i varies over the atoms of ν^+. It follows at once that

$$- 2J(\nu_1, \nu_2) < nc^2/30. \tag{30}$$

We may now conclude from (29) and (30) that

$$\begin{aligned} -I(\nu) > -I(\Phi_1 * \nu) &> (1/8 - 1/30)nc^2 \\ &= .09nc^2. \end{aligned} \tag{31}$$

The inequality (31) implies inequality (28).

Proceeding as in the proof of Theorem 15 $D_s(\nu)^2(2\pi\delta_2) > -I(\nu)$, and $-I(\nu) > -I(\Phi_1 * \nu) > .045\ \bar{\delta}nc^2$. These inequalities at once give inequality (27) and Theorem 17. ∎

Roughly stated, the reciprocals of δ_1, δ_1^+, etc., are analogous to bounds on the various density functions associated with continuous measures.

There are various other ways in which Theorem A might be extended under special hypotheses. One way is to assume that the atoms of ν^- have much larger measure than the atoms of ν^+. This assumption would allow estimates of discrepancy that are independent of $\delta^\#$. The limiting case would have ν^- as a continuous measure. The next section is devoted to a problem of this type.

7. Measures with continuous parts, Theorem B

By varying the hypotheses on ν there are numerous ways in which the previous work on discrete measures might be generalized. We shall concentrate on a class which includes many of the measures that have been previously investigated in connection with problems of discrepancy.

The symbol $\Omega(E^t)$ will denote the class of signed measures ν of total mass zero on E^t such that the positive part ν^+ is Lebesgue t-measure restricted to some subset F of E^t, and the negative part ν^- is an atomic measure concentrated on the points $p_1, p_2, \ldots, p_n$ in E^t.

Lemma 18. *Let ν be a nonzero measure in $\Omega(E^2)$, then*

$$-I(\nu) \geq -I(\Phi_1 * \nu^+) + 1/16 \sum_i (\nu p_i)^2 - 1.58 \sum_i |\nu p_i|. \tag{32}$$

Proof. As before we shall establish the inequality for $-I(\Phi_1 * \nu)$. Again letting $\nu = \nu^+ - \nu^-$, $\nu_1 = \Phi_1 * \nu^+$, and $\nu_2 = \Phi_1 * \nu^-$, and as before $-I(\Phi_1 * \nu) = -I(\nu_1) + 2J(\nu_1, \nu_2) - I(\nu_2)$.

Without additional hypotheses $I(\nu_1) = I(\Phi_1 * \nu^+)$ cannot be bounded away from zero. Nonetheless it should be noted that the formula for $I(\nu_1)$ becomes

$$\begin{aligned} I(\nu_1) &= \iint J(p,q)\,d\nu^+(p)\,d\nu^+(q) \\ &= \iint \chi(p)\chi(q)J(p,q)\,dm(p)\,dm(q), \end{aligned}$$

where m is Lebesgue measure and χ is the characteristic function of the set F which supports ν^+. Most certainly $-I(\nu_1) > 0$.

Because Lemma 16 insures that $J(\Phi_1, d)$ remains negative, as in inequality (29) it can also be stated that

$$-I(\nu_2) \geq -I(\Phi_1) \sum_i (\nu^- p_i)^2 = 1/16 \sum_i (\nu^- p_i)^2.$$

With the present hypotheses on the measure ν the "mixed" term $J(\nu_1, \nu_2)$ may be represented as

$$J(\nu_1, \nu_2) = \sum_i (\nu^- p_i) \int J(p_i, q)\,d\nu^+(q).$$

Since $J(\Phi_1, d)$ is always negative $J(\nu_1, \nu_2)$ is negative. It follows that $-\int J(p_i, q)d\nu^+(q) = -\int J(p_i, q)\chi(q)dm(q) < -\int J(p_i, q)dm(q)$. By converting to polar coordinates $-\int J(p_i, q)dm(q) = -2\pi \int J(\Phi_1, r)r\, dr$, $0 < r < \infty$. Lemmas 16 and 11 allow the estimates concerning $J(\Phi_1, r)$:

$$-\int_{r\leq 2} J(\Phi_1, r)rdr < -I(\Phi_1)\int_{r\leq 2} rdr = 1/8, \tag{33}$$

$$-\int_{r>2} J(\Phi_1, r)rdr < \int_{r>2} (3/4096)r^{-3}dr = 3/8192. \tag{34}$$

Inequalities (33) and (34) together with the previous observations give

$$0 < -\int J(p_i, q)d\nu^+(q) < 1.58/2.$$

Noting that $|\nu p_i| = \nu^- p_i$, inequality (32) follows directly upon combining the various estimates obtained above. This completes the proof of the lemma, and we proceed to the proof of Theorem B. ∎

Let ν^+ be the restriction of Lebesgue measure to a disk of area K, while ν^- assigns measure K/n to each of the n points $p_1, \ldots, p_n$. We shall (falsely) assume that information about $-I(\Phi_1 * \nu^+)$ is absent. Inequality (32) becomes

$$-I(\nu) > 1/16\sum_i (K/n)^2 - 1.58\sum_i K/n = \\ = (1/16)K^2/n - 1.58K.$$

Let us choose $K = 100n$ which means that $\nu^- p_i = 100$ for each i. It quickly follows that

$$-I(\nu) > 400n,$$

and there remains a problem of scaling the present configuration to that required by Theorem B.

The radius of the present disk is $(10/\sqrt{\pi})n^{1/2}$. First, dilate the plane by a similarity τ of ratio $(1/10)n^{-1/2}$ centered at the center of the disk, and then define the measure $\bar{\nu}$ by

$$\bar{\nu}A = (1/100n)\nu(\tau^{-1}A).$$

The scaling properties of the functional I imply that

$$-I(\bar{\nu}) = -(1/100000n^{5/2})I(\nu) > (1/250)n^{-3/2}. \tag{35}$$

Using our standard argument at this point yields

$$\{D_s(\bar{\nu})\}^2 2\sqrt{\pi} > -I(\bar{\nu}). \tag{36}$$

It follows at once from (35) and (36) that

$$D_s(\bar{\nu}) \geq (500\sqrt{\pi})^{-1/2}n^{-3/4} > .0335n^{-3/4}.$$

Theorem B is now immediate.

8. Corollary theorems to the proof of Theorem B

The following theorem is contained in the proof to Theorem B.

Theorem 19. *Let the measure* ν *belong to the class* $\Omega(E^2)$ *with* $mF = 1$ *and* $\nu^-(p_i) = 1/n$ *for each* i, *then*

$$-I(\nu) > .004n^{-3/2}. \tag{37}$$

Proof. Inequality (37) is contained in inequality (35). ∎

Theorem 20. *Let the measure* ν *belong to the class* $\Omega(E^2)$ *with* $mF = 1$ *and* $\nu^-(p_i) = 1/n$ *for each* i. *Suppose that* ν *is supported within a disk of radius* R. *Then*

$$nD_S(\nu) > .0252R^{-1/2}n^{1/4}. \tag{38}$$

Proof. Following along the proof of Theorem B, the measure of the lines that intersect the support of the measure ν will be bounded by $2\pi R$ rather than $2\sqrt{\pi}$ in the special situation of Theorem B. Making this substitution in inequality (36) leads at once to inequality (38). A further sharpening can be obtained by replacing the number $2\pi R$ by the perimeter of the convex hull of the support for ν. ∎

Theorem 21. *Let the measure* ν *belong to the class* $\Omega(E^2)$ *with* $mF = 1$ *and* $\nu^-(p_i) = 1/n$ *for each* i. *Suppose that* ν *is supported within a disk of radius* R, *and let* $\rho < R$ *be a preassigned positive number. Then there is an open square* S *of side* ρ *such that*

$$n\nu(S) > .0252\rho^2\bar{R}^{-5/2}n^{1/4}$$

where $\bar{R} = \rho + \rho[[R/\rho]]$.

Proof. (Sketched) For simplicity suppose that R/ρ is an integer. Consider the collection of grids that divide the plane into squares of side ρ. Fundamental principles from the theory of measure and integration assure that there exists a grid such that every grid line has ν-measure zero, and at least one line of the grid determines a halfplane with ν-measure exceeding $.0252R^{-1/2}n^{-3/4}$. Theorem 20 allows this estimate. In this halfplane it is clear that there is a square of side $\bar{R} = R + \rho$ having sides bounded by grid lines, and having ν-measure exceeding $.0252\bar{R}^{-1/2}n^{-3/4}$. If follows at once that one of the $\bar{R}^2/\rho^2$ grid squares contained in the large square has the required ν-measure. ∎

9. Higher dimensions, Theorem A

Lemma 22. *Let* Φ_r *belong to* $\Psi[-\varepsilon/2, \varepsilon/2]$ *for some* $\varepsilon \leq 1/2$ *and satisfy* $I^{2k}(\Phi_r) = 0$ *for* $1 \leq k \leq r$. *Furthermore, suppose that* $\|\Phi_r\| = 1$. *Then if* $d \geq 2$, *the following inequality holds.*

$$|J(\Phi_r, d)| \leq 1.07\varepsilon^{2r+2}/d^{2r+1}. \tag{39}$$

Proof. From equation (16) it follows that

$$|J(\Phi_r, d)| \leq \sum_{k \geq r+1} |c_k I^{2k}(\Phi_r)| d^{-2k+1}.$$

The facts that $|c_k| < 1$ and $|I^{2k}(\Phi_r)| \leq \varepsilon^{2k}$ allow one to write $|J(\Phi_r, d)| < d(\varepsilon^2/d^2)^{r+1}\{1 - \varepsilon^2/d^2\}^{-1}$. Since $\varepsilon^2/d^2 \leq 1/16$, the inequality (39) follows. ∎

Lemma 23. *Let the n points lie p_i in E^t, let $\delta_1 = \min |p_i - p_j| \geq 2$, and let Φ_r satisfy the hypotheses of Lemma 22 where $t \leq 2r$. Then if the measure ν is supported by the p_i,*

$$\sum_{i \neq j} |J(p_i, p_j)| \nu p_i \nu p_j \leq 1.07(2\varepsilon)^{t+2}(t+1) \sum_i (\nu p_i)^2. \tag{40}$$

Proof. Fix one of the p_i and call this point p_0 and let d_i, $1 \leq i \leq n-1$, be the distances from p_0 to the other points, indexed in order of increasing magnitude. Then

$$d_k \geq k^{1/t}/2. \tag{41}$$

To see inequality (41) note that the t-disk of radius $d_k + 1$ centered at p_0 must contain $k+1$ pairwise disjoint unit t-disks. It follows that $k + 1 \leq (d_k + 1)^t$.

Applying inequalities (39) and (41) and summing gives

$$\begin{aligned} \sum_k |J(\Phi_r, \Phi_k)| &\leq 1.07\varepsilon^{2r+2} \sum_k 1/d_k^{2r+1} \\ &\leq 1.07\varepsilon^{t+2} \sum_k 1/d_k^{t+1} \\ &\leq 1.07\varepsilon^{t+2} 2^{t+1} \sum_k k^{-(1+1/t)} \\ &\leq 1.07(2\varepsilon)^{t+2}(t+1). \end{aligned} \tag{42}$$

Finally, note that the estimate (42) is independent of the choice of any of the n points p_i. Inequality (40) follows at once from Lemma 12. ∎

Let us introduce notation as follows. By Theorem 8 there is an atomic measure $\Phi(r)$ of unit norm, supported by at most $2r + 2$ points x_i in the interval $[-1/2, 1/2]$ such that for $1 \leq k \leq r$, $I^{2k}[\Phi(r)] = 0$. For $\varepsilon \leq 1/2$, let $\Phi(r, \varepsilon)$ be the dilated measure defined by $\Phi(r, \varepsilon)(x_i) = \Phi(r)(x_i/\varepsilon)$. Observe that $I^\alpha[\Phi(r, \varepsilon)] = \varepsilon^\alpha I^\alpha[\Phi(r)]$ and that $\|\Phi(r, \varepsilon)\| = \|\Phi(r)\| = 1$.

Theorem 24. *Let ν be a measure of total mass 0 concentrated on the points p_i in E^t where $t \leq 2r$. Then*

$$|I(\nu)| \geq (\delta_1/2)\{\varepsilon |I[\Phi(r)]| - 1.07(2\varepsilon)^{t+2}(t+1)\} \sum_i (\nu p_i)^2. \tag{43}$$

Proof. Let us assume that $\delta_1 = 2$. The result for arbitrary δ_1 follows from the linearity of the functional I with respect to dilations of the measure ν. Following the usual patterns note that $-I[\Phi(r, \varepsilon) * \nu] \leq -I(\nu)$ and that by (14)

$$I[\Phi(r, \varepsilon) * \nu] = \sum_{i \neq j} J(p_i, p_j) \nu p_i \nu p_j + I[\Phi(r, \varepsilon)] \sum_i (\nu p_i)^2.$$

By the triangle inequality

$$(44)\qquad |I[\Phi(r,\varepsilon)*\nu]| \geq |I[\Phi(r,\varepsilon)]\sum_i(\nu p_i)^2| - \Big|\sum_{i\neq j} J(p_i,p_j)\nu p_i \nu p_j\Big|$$

The inequality (44) is strengthened by replacing $\sum_{i\neq j} J(p_i,p_j)\nu p_i\nu p_j$ by the right side of inequality (40). The fact that $I[\Phi(r,\varepsilon)] = \varepsilon I[\Phi(r)]$ leads at once to the inequality (43). ∎

Remark. For any choice of $\Phi(r)$ and t, $\varepsilon|I[\Phi(r)]| - 1.07(2\varepsilon)^{t+2}(t+1)$ is a positive number for sufficiently small values of ε. It is quite possible that $\Phi(r)$ can always be chosen so that $\varepsilon = 1/2$ will be satisfactory.

Remark. In the present work the measure μ on the hyperplanesets of E^t are normalized so that $|p-q| = 1/2\mu\{h : h \text{ cuts seg } pq\}$. In a more usual normalization $\mu\{h : h \text{ cuts seg } pq\} = (t-1)^{-1}O_{t-2}|p-q|$ where O_k is the measure of the unit spherical surface S^k. (See [10] p.229 13.71 and p.233 14.2.) The present normalization is more convenient for the study of the functional I.

Proof of Theorem A. The μ-measure of the set of hyperplanes that cut the convex hull of a set of diameter δ_2 is no greater than the μ-measure of the set of hyperplanes that cut a $t-1$ sphere of diameter δ_2. With the standard normalization this measure equals $(1/2)O_{t-1}\delta_2$. (See [10] p.233, 14.2.) It follows that with the present normalization the μ-measure of the set of hyperplanes that cut the convex hull of a set of diameter δ_2 is no greater than $(t-1)O_{t-1}(O_{t-2})^{-1}\delta_2$. Combining these observations with the inequality (43) gives the inequality

$$\{D_s(\nu)\}^2\{(t-1)O_{t-1}(O_{t-2})^{-1}\delta_2\} \geq (\delta_1/2)\{\varepsilon|I[\Phi(r)]| - 1.07(2\varepsilon)^{t+2}(t+1)\}\sum_i(\nu p_i)^2.$$

By choosing $\Phi(s), t \leq 2s$, and $\varepsilon > 0$ sufficiently small so that the number $\varepsilon|I[\Phi(r)]| - 1.07(2\varepsilon)^{t+2}(t+1)$ is positive, it is clear that the inequality (1) holds where

$$2c_t^2 = \{\varepsilon|I[\Phi(r)]| - 1.07(2\varepsilon)^{t+2}(t+1)\}\{(t-1)O_{t-1}(O_{t-2})^{-1}\}^{-1}.$$

This completes the proof of Theorem A. ∎

In E^t if ν^- assigns measure $1/n$ to each of the n points p_i and ν^+ is Lebesgue t-measure restricted to a t-disk of unit volume, a direct generalization of the proof of Theorem B would lead to the inequality

$$(45)\qquad nD_s(\nu) > c_t n^{1/2(1-1/t)}.$$

However, there is an unanswered question that leaves a gap in the proof. The question is whether a suitable measure $\Phi(r)$, $t \leq 2r$, can be found that satisfies the conclusion of Lemma 16. Bruce Reznick has given a pretty argument that the measure Φ_2 described in Example 2 above does satisfy the conclusions of Lemma 16. In particular, $-J(\Phi_2, d) > 0$ for all d and $J'(\Phi_2, d) > 0$. This means that all of the techniques of this paper work in E^t for $t \leq 4$, and generally on embedded t-manifolds of dimensions $t \leq 4$. An affirmative solution of the following problem would remove this bound on t.

Problem 1. Generalize Lemma 16 and Reznick's result by showing that for all positive integers r a suitable measure $\Phi(r)$ can be found that satisfies $J'(\Phi(r), d) > 0$ for all d. Can $J(\Phi, d) = 0$ ever for a Φ in Ψ ?

There are a number of weaker results that, once established, would allow the full method to extend into all dimensions.

10. Discrepancy on spheres, Stolarsky's formula

In this section we give a brief outline of the present method in the case of S^2 the unit 2-sphere. The application of the method is not overly sensitive to the geometry of the support for a measure beyond assumptions concerning dimension and smoothness. Therefore the treatment of S^2 is very similar to the previous treatment of the Roth disk problem. Only constants change. The general 2-sphere of arbitrary radius will be denoted by $\bar{S}^2$.

The sphere $\bar{S}^2$ is assumed to lie in the space E^3 so that signed Borel measures on $\bar{S}^2$ are treated as signed measures on E^3. The measure μ is defined on 3-planesets and has the normalization given by (3) above. All distances considered are euclidean.

Let ν^+ assign measure K/n to each of n points p_i on the sphere $\bar{S}^2$ of area K, and let ν^- be the usual 2-measure on $\bar{S}^2$. Now $I(\nu^+ - \nu^-) = I(\nu^+) - 2J(\nu^+, \nu^-) + I(\nu^-)$. However, as a result of the symmetry of the measure ν^- it is seen that $J(\nu^+, \nu^-) = I(\nu^-) = (2/3)\pi^{-1/2}K^{3/2}$. Combining this with the fundamental equation (6a) gives

$$I(\nu^+) + \int A^2(h)d\mu(h) = (2/3)\pi^{-1/2}K^{3/2}. \tag{46}$$

Formula (46) for dimension 2 expresses in the context of the present paper the pretty relation discovered by Stolarsky [15], who then used estimates by Schmidt [13] on the discrepancy of spherical caps to estimate $I(\nu^+)$. This paper of Stolarsky is the first place where the close relation between discrepancy and distance sums is revealed. The ideas of integral geometry, as expressed by formula (5), along with the inner product $-J$ allow a pleasing derivation of the Stolarsky formula. Of course, the same ideas give a Stolarsky formula for spheres in all dimensions. The Stolarsky formula may actually characterize spheres in that the author knows of no other situation where maximizing $I(\nu^+)$ is equivalent to minimizing $\int A^2$, at least for ν in the class Ω.

To apply the methods of this paper to S^2 (in brief outline) consider the measure $\nu * \Phi_1$ as a measure on E^4 in the established manner. The measures ν_1 and ν_2 are defined as before, it is necessary to estimate

$$I(\Phi_1 * \nu) = I(\nu_1) - 2J(\nu_1, \nu_2) + I(\nu_2).$$

Next one establishes the inequality

$$|2J(\nu_1, \nu_2)| < c_1 K \tag{47}$$

where c_1 is a positive constant independent of K. Of course, the constant c_1 is analogous to the number 1.58 appearing in (32). Also

$$-I(\nu) > -(K/n)^2\Big\{\sum_{ij} J(p_i, p_j) + I(\Phi_1)\Big\} > -(K^2/n)I(\Phi_1).$$

Lemma 16 allows the last inequality. Since $-I(\nu_2) > 0$, one can now write

$$-I(\nu_1) = -I(\Phi_1 * \nu) > -I(\Phi_1)K^2/n - c_1 K = (1/16)K^2/n - c_1 K.$$

Letting $K = c_2 n$ for a rather large choice of c_2 yields $-I(\nu) > -c_3 n$. The radius of the associated sphere is $c_4\sqrt{n}$. It then becomes a simple task of scaling to the case $K = 4\pi$ to obtain for S^2

$$nD_s(\nu) > c_5 n^{1/4}. \tag{48}$$

The inequality (48) was first obtained by Beck [6] along with the higher dimensional versions corresponding to (45) above. However we state a more general theorem for dimension 2 analogous to Theorem 20 above. The proof is a straightforward modification of the foregoing discussion. An important point here is that inequality (47) depends only on very crude properties of the measure ν.

Theorem 25. *Let $\bar{S}^2$ be a sphere of radius R containing the n points p_i. Let ν^- be Lebesgue surface measure restricted to a set of measure* 1, *and ν^+ assign measure $1/n$ to each p_i. Then*

$$nD_s(\nu) > c_2 R^{-1/2} n^{1/4}.$$

11. Remarks concerning bounds for I

It is clear that the methods of this article are actually based on finding estimates for the functional $I(\nu)$ rather than the elusive $D_s(\nu)$. It is only by way of a remarkable estimate by Beck [7,p.180] that one can get an idea of just how small $D_s(\nu)$ can be in the case of the Roth disk segment problem. Beck shows that for each n there is a signed measure ν_n, where $\nu_n^+(p_i) = 1/n$ on the n points p_i and ν_n^- is Lebesgue area measure on a disk of unit area, that satisfies the inequality

$$nD_s(\nu_n) < cn^{1/4}(\log n)^{1/2}. \qquad (J.Beck)$$

Thanks to this fine result, it is clear that the lower bound estimates on D_s via I in Theorem B are almost unreasonably good. The following problem is probably difficult.

Problem 2. Is it true for any such sequence that $\lim n^{3/4} D_s(\nu_n) = \infty$?

Returning to the functional I one might ask how accurate is the estimate of Theorem 19 that $-n^2 I(\nu) > .004n^{1/2}$ for the n point disk segment problem. In 1972 the author noted a rather easy way to obtain upper bounds on how small I could become. Again, the surprise is how accurate the estimates are. The method was first used to obtain upper bounds for distance sums on 2-spheres [1]. We illustrate the method on the measure of the disk segment problem.

A planar disk K of area n^2 if the union of n^2 pairwise disjoint sets A_i if area 1, each having diameter less than $4\sqrt{\pi}$. For each i let p_i be chosen at random with respect to Lebesgue measure in A_i. Define the measure ν by $\nu^+(p_i) = 1$ and ν^- as Lebesgue measure on K. Then is not difficult to show that the expectation $EI(\nu)$ satisfies the relation

$$-EI(\nu) = \sum -EI(\nu_i) < n^2 8/\sqrt{\pi} \tag{49}$$

where ν_i is the measure in $\Psi(E^2)$ obtained by restricting ν to A_i. the following lemma is the key.

Lemma 26. *Let A_1 and A_2 be sets in E^t of unit Lebesgue t-measure. Let P_i be a random point in A_i, $i = 1, 2$. Define the measure ν_i in $\Psi(E^t)$ by setting $\nu_i^+ p_i = 1$ and ν_i^- to be Lebesgue measure restricted to A_i. Then*

$$EJ(\nu_1, \nu_2) = 0.$$

Proof. Note that $EJ(\nu_1^+, \nu_2^+) = E|p_1 - p_2| = \int\int |p_1 - p_2| d\nu_1^- d\nu_2^- = J(\nu_1^-, \nu_2^-)$, and that $EJ(\nu_1^+, \nu_2^-) = E \int |p_1 - p_2| d\nu_2^- = \int\int |p_1 - p_2| d\nu_1^- d\nu_2^- = J(\nu_1^-, \nu_2^-)$. Certainly $EJ(\nu_1^-, \nu_2^-) = J(\nu_1^-, \nu_2^-)$. Thus $EJ(\nu_1, \nu_2) = EJ(\nu_1^+ - \nu_1^-, \nu_2^+ - \nu_2^-)$ which is the sum of four numbers $\pm J(\nu_1^-, \nu_2^-)$ that cancel one another. This proves the lemma. ∎

To establish (49), $EI(\nu) = \sum_i EI(\nu_i) + \sum_{i \neq j} EJ(\nu_i, \nu_j) = \sum_i EI(\nu_i)$. It is quickly seen that $-I(\nu_i) \leq 2$(diameter A_i) for any choice of p_i. Therefore it is clear that there is a measure ν of this class such that $-I(\nu) < n^2 4 \sqrt{\pi}$. Scaling leads at once to an upper bound theorem.

Theorem 27. *For each integer n there are n points p_i in a disk of unit area such that for the signed measure ν_n of the disk segment problem*

$$-n^2 I(\nu_n) < (8/\sqrt{\pi}) n^{1/2}.$$

Theorem 27 shows that $n^{-3/2}$ is the optimal order of magnitude for the sequence $\sup |I(\nu_n)|$ for the measures of the disk segment problem. It is clear that the idea behind this estimate has generalizations to many other functionals and in all dimensions.

12. Final problems and remarks

Problem 3. Let ν be a measure in $\Omega(E^t)$ where $\nu^- p_i = 1/n$ and ν^+ is Lebesgue measure restricted to a set F of unit t-measure. If n remains fixed while the points p_i and the set F are allowed to vary, determine the infimum of the positive numbers $-I(\nu)$.

For the E^t, a collection of widely spaced disks of volume $1/n$ centered at the p_i, no $t+1$ of the p_i lying in a hyperplane, probably approaches optimality. The number $t/2n$ gives the attainable minimum for the number $D_s(\nu)$. This is one example of a problem where $D_s(\nu)$ is easier to analyze than is $I(\nu)$.

Problem 4. To what extent is Stolarsky's identity characteristic of the Euclidean sphere ?

Part of Problem 4 lies in formulating a precise statement. In [17] Stolarsky studies a class of metric spaces where his identity does generalize. Also, a number of problems related to distributions of points and the functional I relative to Euclidean spheres have been considered in the papers [4], [16]. Perhaps recent estimates will allow further progress.

Problem 5. Let $2n$ points p_i lie in a Hilbert space with $|\nu p_i| = 1$ and $\sum \nu p_i = 0$. Is there an absolute constant c such that $-I(\nu) \geq c\delta_1 n$ where δ_1 is the minimum distance between two p_i ?

A resolution of Problem 5 would be an important result in the study of metric inequalities. If the resolution were negative, then there would be a need for an asymptotic estimate of the best dimensional constants c_t. The methods of this paper produce a c_t, but there is no estimate of optimal magnitude. Perhaps letting the p_i be the set of vertices of a unit t-cube provides an interesting example for study.

Presently, an article studying the application of some of the present methods to the functionals I^α and J^α for $0 < \alpha \leq 2$ is in preparation. This range of α has been important to the study of metric embedding theory as well as to the study of related topics such as *transfinite diameter*, (See [5]). Also, Mr. Allen D. Rogers is investigating the method as it applies to signed measures ν with continuous parts ν^+, ν^-, as well as applications to norm inequalities for certain related transformations, such as Radon transformations. The author wishes to express his gratitude to Mr. Rogers for his assistance in proofreading the article.

Addendum

The author has recently settled the key issue in Problem 1. We state the following theorem without proof.

Theorem. *Let Φ be any nonzero atomic measure in the class $\Psi(\mathbb{R})$. Then $-J(\Phi, d)$ is a monotone decreasing positive function.*

This theorem allows the extension of the key ideas of the present paper to any E^t for measures ν with nonatomic parts. However, a proof requires a nontrivial extension of several ideas relating to Lemma 16. A complete discussion of this theorem will appear in the article under preparation.

References

[1] R. ALEXANDER: On the sum of distances between n points on a sphere, *Acta Math. Acad. Sci. Hungar.*, **23** (1972), 443–448.

[2] R. ALEXANDER: Generalized sums of distances, *Pacific Jour. Math.*, **56** (1975), 297–304.

[3] R. ALEXANDER: On the sum of distances between n points on a sphere. II, *Acta Math. Acad. Sci. Hungar.*, **29** (1977), 317–320.

[4] R. ALEXANDER: Metric averaging in Euclidean and Hilbert spaces, *Pacific Jour. Math.*, **85** (1979), 1–9.

[5] R. ALEXANDER, and K. B. STOLARSKY: Extremal problems of distance geometry related to energy integrals, *Trans. Amer. Soc.*, **193** (1973), 1–31.

[6] J. BECK: Sums of distances between points on a sphere — an application of the theory of irregularities of distribution to discrete geometry, *Mathematica,* **31** (1984), 33–41.

[7] J. BECK, and W. W. L. CHEN: *Irregularities of distribution,* Cambridge Tracts in Mathematics, **89**, Cambridge University Press, Cambridge, 1987.

[8] L. KUIPERS, and H. NIEDERREITER: *Uniform distribution of sequences,* John Wiley, New York, **1974**.

[9] K. F. ROTH: On irregularities of distribution, *Mathematica,* **1** (1954), 73–79.

[10] L. A. SANTALÓ: *Integral geometry and geometric probability,* Encyclopedia of Mathematics and its Applications, **1**, Addison–Wesley, Reading, Mass., 1976.

[11] 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.

[12] W. M. SCHMIDT: Irregularities of distribution. IV, *Invent. Math.*, **7** (1969), 55–82.

[13] W. M. SCHMIDT: Irregularities of distribution. VII, *Acta Arith.*, **21** (1972), 45–50.

[14] K. B. STOLARSKY: Sums of distances between points on a sphere, *Proc. Amer. Math Soc.*, **35** (1972), 547–549.

[15] K. B. STOLARSKY: Sums of distances between points on a sphere. II, *Proc. Amer. Math. Soc.*, **41** (1973), 575–582.

[16] K. B. STOLARSKY: Spherical distributions of n points with maximal distance sums are well spaced, *Proc. Amer. Math. Soc.*, **48** (1975), 203–206.

[17] K. B. STOLARSKY: *Discrepancy and sums of distances between points of a metric space,* The geometry of metric and linear spaces, Ed. L. M. Kelly, Springer–Verlag, **1975**, 44–55.

[18] H. WEYL: Über die Gleichverteilung von Zahlen mod. Eins, *Math. Ann.*, **77** (1916), 313–352.

Ralph Alexander

Department of Mathematics,
University of Illinois,
Altgeld Hall, 1409 W. Green St.
Urbana, Illinois 61801
U. S. A.