NUMERICAL GEOMETRY—NUMBERS FOR SHAPES

JOAN CLEARY AND SIDNEY A. MORRIS

Mathematics Department, La Trobe University, Bundoora, Victoria, Australia, 3083

DAVID YOST

Mathematics Department, I.A.S., Australian National University, G.P.O. Box 4, Canberra, A.C.T. Australia, 2601

1. Introduction. In 1964, O. Gross [9] published a short but intriguing paper. His main result remains little known and is, in his and our experience, greeted with a range of reactions from surprise to disbelief. Before stating the result we give some concrete examples:

Let Σ^1 denote the circle (not the disk) of unit diameter and x_1, x_2, \ldots, x_n any points on Σ^1 . (Throughout, sets of unit diameter will be more convenient to work with than sets of unit radius.) Then there is a point y on Σ^1 such that the average distance from y to x_1, x_2, \ldots, x_n is $2/\pi$. The number $2/\pi$ works for all collections of n points, for any positive integer n. Moreover, no other number will work! For the sphere Σ^2 of unit diameter in Euclidean 3-space, a similar result is true except that $2/\pi$ is replaced by 2/3. For Σ^3 and Σ^4 the "magic numbers" are $32/15\pi$ and 72/105, respectively. An equilateral triangle with sides of length one has "magic number" $(2 + \sqrt{3})/6$, while a semicircle of unit diameter has "magic number" $4/(4 + \pi)$.

Without further ado we state the Gross theorem.

THEOREM 1 [GROSS, 9]. If (X, d) is any compact connected metric space then there is a unique positive real number a(X, d) with the following property: for each positive integer n and for all (not necessarily distinct) x_1, x_2, \ldots, x_n in X, there exists a y in X such that $\frac{1}{n}\sum_{i=1}^n d(x_i, y) = a(X, d)$.

Several questions immediately present themselves:

- (a) Given (X, d), how does one find a(X, d)?
- (b) What does the number a(X, d) tell us about (X, d)?
- (c) What values can a(X, d) take on?
- (d) How does one prove the Gross theorem?
- (e) How might the Gross theorem be applied to other areas of mathematics?

The first of these questions is the most tantalizing.

2. Some elementary examples. In this section we show how to calculate a(X, d) in some simple cases. Our first example shows that if (X, d) is the unit interval [0, 1] with the usual metric,

Joan Cleary did her B.Sc (Hons) in Pure Mathematics at La Trobe University in 1982 with her honours project being on Numerical Geometry. She is currently a high school math teacher.

Sidney Morris is a graduate of the University of Queensland and The Flinders University of South Australia, obtaining his Ph.D. in 1970. He has held various positions at University of Queensland, University of Adelaide, University of Florida, and University of New South Wales, and Visiting Professorships at Tulane University and Tel-Aviv University. In 1974 he was a Science Research Council Visiting Fellow at the University College of North Wales. Since 1976 he has been a faculty member of La Trobe University. For $5\frac{1}{2}$ years he was Editor of the Bulletin of the Australian Mathematical Society and is currently Editor-in-Chief of the "Australian Mathematical Society Lecture Series"—a new series of books published by Cambridge University Press. He is the author of about 80 papers, most of which are on topological groups, and the book "Pontryagin Duality and the Structure of Locally Compact Abelian Groups."

David Yost is a graduate of The University of Melbourne, The Australian National University, and Edinburgh University, with his Ph.D. being conferred in 1980. He is currently a Research Fellow at The Australian National University. He has held a Queen Elizabeth II Fellowship and a La Trobe University Research Fellowship. Most of his publications are in Banach space theory.

then a(X, d) exists, is unique, and equals 1/2.

EXAMPLE 1 [11]. Let X = [0,1] and d be the usual metric on X. We show firstly that a(X,d) exists. {To do this it helps to have a peep at the answer. If a(X,d) exists, then, putting n=2, $x_1=0$, and $x_2=1$, we obtain for all y in $X=[0,1], \frac{1}{2}\sum_{i=1}^2 d(x_i,y)=\frac{1}{2}[(y-0)+(1-y)]=\frac{1}{2}$. So a(X,d) would be unique and equal $\frac{1}{2}$.}

To see that a(X, d) exists we proceed as follows: let $x_1, x_2, ..., x_n$ be any points in X and consider the continuous "average distance" function $f: [0, 1] \to [0, 1]$ given by

$$f(t) = \frac{1}{n} \sum_{i=1}^{n} |x_i - t|, \quad t \in [0,1].$$

Clearly

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

So either $f(0) \le 1/2 \le f(1)$ or $f(1) \le 1/2 \le f(0)$. Applying the Intermediate Value Theorem yields the existence of $y \in [0,1]$ such that f(y) = 1/2.

So we have shown that for any positive integer n and any x_1, x_2, \ldots, x_n in [0, 1], there exists a point y in [0, 1] such that $\frac{1}{n} \sum_{i=1}^{n} d(x_i, y) = 1/2$. Thus 1/2 is seen to satisfy the conditions for a(X, d) in Theorem 1. So existence of a(X, d) is verified, and uniqueness follows from the remarks above in braces.

In the next three examples we assume the existence, but not the uniqueness, of a(X, d). The technique used in Examples 2 and 3 should ring a bell for those familiar with game theory and, in particular, with the mini-max theorem. We shall say more about mini-max methods in §5.

EXAMPLE 2 [5]. Let X be the closed ball with centre O and radius 1/2 in Euclidean m-space, \mathbb{R}^m , and let d be the Euclidean metric on \mathbb{R}^m .

Assume a(X, d) exists. Firstly, let n = 1 and $x_1 = 0$. Then for any y in X, $d(x_1, y) \le 1/2$, and so $a(X, d) \le 1/2$.

Next, put n=2 and choose diametrically opposite boundary points x_1' and x_2' . Then $d(x_1', x_2') = 1$, and for any y in X, $\frac{1}{2}[d(x_1', y) + d(x_2', y)] \ge \frac{1}{2}d(x_1', x_2') = \frac{1}{2}$. So $a(X, d) \ge \frac{1}{2}$. Thus a(X, d) must equal $\frac{1}{2}$ (and so is unique).

EXAMPLE 3 [5]. Let X be the equilateral triangle (not its convex hull, the triangular region) with sides of length one and d the Euclidean metric. We show here that if a(X,d) exists then it must equal $\frac{2+\sqrt{3}}{6}$.

Firstly, let n = 3 and x_1, x_2 , and x_3 be the vertices of the triangle. (See Fig. 1.)

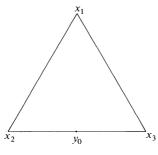


Fig. 1

It is easily verified that for $y \in X$, the quantity

$$d(x_1, y) + d(x_2, y) + d(x_3, y)$$

achieves its minimum value when y is the midpoint y_0 of any one of the sides. So for all y in X,

$$\frac{1}{3}\sum_{i=1}^{3}d(x_i,y)\geqslant \frac{1}{3}\sum_{i=1}^{3}d(x_i,y_0)=\frac{1}{3}\left(\frac{1}{2}+\frac{1}{2}+\frac{\sqrt{3}}{2}\right)=\frac{2+\sqrt{3}}{6}.$$

Thus $a(X, d) \ge \frac{2 + \sqrt{3}}{6}$.

We now proceed to show that $a(X, d) \le \frac{2 + \sqrt{3}}{6}$. Again, put n = 3 and let x_1' , x_2' , and x_3' be the midpoints of the sides. (See Fig. 2.)

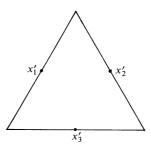


Fig. 2

The quantity $d(x'_1, y) + d(x'_2, y) + d(x'_3, y)$ achieves its maximum value when y is any vertex y_1 . So for all y in X,

$$\frac{1}{3}\sum_{i=1}^{3}d(x_{i}',y)\leqslant \frac{1}{3}\sum_{i=1}^{3}d(x_{i}',y_{1})=\frac{1}{3}\left(\frac{1}{2}+\frac{1}{2}+\frac{\sqrt{3}}{2}\right)=\frac{2+\sqrt{3}}{6}.$$

Thus $a(X, d) \leq \frac{2 + \sqrt{3}}{6}$.

Consequently $a(X, d) = \frac{2 + \sqrt{3}}{6}$.

We point out that the same method can be used to calculate the magic number of each regular polygon in \mathbb{R}^2 . This yields: If X_n is a regular n-gon of diameter one, then

$$a(X_n, d) = \frac{1}{2n} \sum_{k=0}^{n-1} \left[\frac{3}{2} + \frac{1}{2} \cos \frac{2\pi}{n} - \cos \frac{2k\pi}{n} - \cos \frac{2(k-1)\pi}{n} \right]^{1/2}$$
, when *n* is even,

$$a(X_n, d) = \frac{1}{n} \sum_{k=0}^{n-1} \left[\frac{\frac{3}{2} + \frac{1}{2} \cos \frac{2\pi}{n} - \cos \frac{2k\pi}{n} - \cos \frac{2(k-1)\pi}{n}}{2 - 2 \cos \frac{(n-1)\pi}{n}} \right]^{1/2}, \text{ when } n \text{ is odd.}$$

(Full details are given in [5].) ■

OPEN QUESTION 1. If X is a general polygon in Euclidean 2-space, what is a(X, d)?

Incidentally, it is easy to see [5] that if X is the perimeter of a rectangle with sides of length l and w, where $l \ge w$, then

$$a(X,d) = \frac{1}{4} \left(l + \sqrt{4w^2 + l^2} \right) / \sqrt{l^2 + w^2}.$$

Example 4 [11]. Let Σ^1 be the circle of centre 0 and radius $\frac{1}{2}$ and d the Euclidean metric. Again, assume $a(\Sigma^1, d)$ exists. Let $x_0, x_1, \ldots, x_{n-1}$ be the points given by

$$x_j = \left(\frac{1}{2}\cos\frac{2\pi j}{n}, \frac{1}{2}\sin\frac{2\pi j}{n}\right), \quad j = 0, 1, \dots, n - 1,$$

which are uniformly distributed on Σ^1 . Then

$$a(\Sigma^{1}, d) = \frac{1}{n} \sum_{j=0}^{n-1} d(x_{j}, y)$$

for some point y on Σ^1 . By symmetry we can choose y to lie on the arc joining x_0 and x_1 . So

$$d(x_0, y) \le d(x_0, x_1) < \sin \frac{\pi}{n} < \frac{\pi}{n}$$

Thus

$$\left| a(\Sigma^{1}, d) - \frac{1}{n} \sum_{j=0}^{n-1} d(x_{j}, x_{0}) \right| < \frac{\pi}{n}.$$

Now

$$d(x_j, x_0) = \sin \frac{\pi j}{n}.$$

Therefore

$$a(\Sigma^{1}, d) = \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \sin \frac{\pi j}{n} = \int_{0}^{1} \sin(\pi x) dx = \frac{2}{\pi}.$$

3. Existence of a(X, d). In Example 2, 3, and 4 we assumed the existence of a(X, d) and proceeded to evaluate it (en route, showing its uniqueness). We give here an elementary proof of the existence of a(X, d). Unfortunately, there is no elementary proof of uniqueness, and so we leave the uniqueness proof to §5.

Proof of Existence [17]. Let (X, d) be a compact connected metric space and put $\mathscr{F} = \bigcup_{n=1}^{\infty} X^n$. Thus \mathscr{F} is the union of all ordered *n*-tuples with members from X. If $F \in \mathscr{F}$ is (y_1, y_2, \ldots, y_n) for some integer n, we define

$$\Theta_F(x) = \frac{1}{n} \sum_{i=1}^n d(x, y_i).$$

[In later sections it will be more convenient to write $\Theta(x, F)$ for $\Theta_F(x)$.] As Θ_F : $(X, d) \to \mathbf{R}^+$ is continuous, $\Theta_F(X)$ is a compact connected subset of \mathbf{R}^+ . So it is a closed bounded interval; that is, $\Theta_F(X) = [a_F, b_F]$ for some $a_F, b_F \in \mathbf{R}^+$.

It is readily seen that a(X,d) exists if and only if $\bigcap_{F \in \mathscr{F}} [a_F,b_F] \neq \varnothing$. We show this by verifying that for all $F, G \in \mathscr{F}, a_F \leqslant b_G$.

Let $G = (z_1, z_2, \dots, z_m)$. Observe that

$$a_F \le \frac{1}{n} \sum_{i=1}^{n} d(z_i, y_i), \text{ for each } l = 1, 2, ..., m,$$

and

$$\frac{1}{m} \sum_{j=1}^{m} d(y_k, z_j) \le b_G$$
, for each $k = 1, 2, ..., n$;

it suffices to show that

$$\frac{1}{n}\sum_{i=1}^n d(z_i, y_i) \leqslant \frac{1}{m}\sum_{j=1}^m d(y_k, z_j)$$

for some $k \in \{1, 2, ..., n\}$ and $l \in \{1, 2, ..., m\}$.

Suppose that for all k = 1, 2, ..., n and l = 1, 2, ..., m

$$\frac{1}{n}\sum_{i=1}^{n}d(z_{i},y_{i})>\frac{1}{m}\sum_{j=1}^{m}d(y_{k},z_{j}),$$

and sum each side over l = 1, 2, ..., m to obtain

$$\sum_{l=1}^{m} \frac{1}{n} \sum_{i=1}^{n} d(z_{l}, y_{i}) > \sum_{j=1}^{m} d(y_{k}, z_{j}).$$

Now summing both sides over k = 1, 2, ..., n we obtain

$$\sum_{l=1}^{m} \sum_{i=1}^{n} d(z_{l}, y_{i}) > \sum_{k=1}^{n} \sum_{j=1}^{m} d(y_{k}, z_{j}).$$

But this is impossible since d is symmetric, so our supposition is false. Hence $a_F \le b_G$ for all F and G in \mathcal{F} . Thus a(X, d) exists.

4. Elton's and Stadje's generalizations. In 1981 Wolfgang Stadje [13] published a more general result than that of Gross, although he was apparently unaware of the Gross Theorem.

THEOREM 2 [STADJE, 13]. If X is any compact connected Hausdorff space and $f: X \times X \to \mathbf{R}$ is a real-valued continuous symmetric function, then there is a unique real number a(X, f) with the following property: for each positive integer n and for all x_1, x_2, \ldots, x_n in X, there exists a y in X such that

$$\frac{1}{n} \sum_{i=1}^{n} f(x_i, y) = a(X, f).$$

REMARK 1. The Stadje Theorem is more general than the Gross theorem since the metric d has been replaced by a continuous symmetric function f. Stadje's Theorem adds an extra "dimension" even for metric spaces, (X, d), since we can put $f = d^2$ or d^3 , etc. (By d^n we mean the function given by $d^n(x, y) = (d(x, y))^n$. For sufficiently large n, d^n will not be a metric.) In fact, as we shall see later Wilson [16] has shown that if (X, d) is a subspace of \mathbb{R}^2 , then it is easier to calculate $a(X, d^2)$ than a(X, d). We note, incidentally, that the proof given in Gross [9] of his theorem easily extends to prove Stadje's Theorem. The elementary existence proof given in §3 also extends to prove the existence of a(X, f).

DEFINITION. If X is a compact connected Hausdorff space and $f: X \times X \to \mathbf{R}$ is a real-valued continuous symmetric function, then we put $D(X, f) = \max\{|f(x, y)|: x, y \in X\}$.

NOTATIONAL CONVENTION. From now on X will be assumed to be infinite and f will be assumed to be not the constant function zero. So D(X, f) > 0. (When f = d is a metric, the second assumption alone implies that X is not a singleton and so, by connectedness, X must be infinite.)

We define the magic number or dispersion number, m(X, f) to be a(X, f)/D(X, f). The next result is a straightforward extension of a result of Gross [9].

PROPOSITION 1 [5]. If (X, d) is any compact, connected metric space and n is any natural number, then $2^{-n} \le m(X, d^n) < 1$.

Proof. Let x_1 and x_2 be points of X such that $d(x_1, x_2) = D(X, d)$. Then Theorem 2 gives us a $y \in X$ with

$$a(X, d^{n}) = \frac{1}{2} (d^{n}(x_{1}, y) + d^{n}(x_{2}, y))$$

$$\ge \left(\frac{1}{2}d(x_{1}, y) + \frac{1}{2}d(x_{2}, y)\right)^{n}$$

$$\ge \left(\frac{1}{2}d(x_{1}, x_{2})\right)^{n} = 2^{-n}D(X, d^{n}).$$

Hence $m(X, d^{-n}) \ge 2^{-n}$.

That $m(X, d^n) \le 1$ is clear, as $a(X, d^n)$ is an average of numbers less than or equal to $D(X, d^m)$. It remains only to show that $m(X, d^n) \ne 1$.

Suppose that $a(X, d^n) = D(X, d^n)$. Again let x_1 and x_2 be diametral points of X. Theorem 2 gives us x_3 in X such that

$$D(X, d^m) = a(X, d^n) = \frac{1}{2} (d^n(x_1, x_3) + d^n(x_2, x_3)),$$

which is the average of two numbers less than or equal to $D(X, d^n)$. Thus we must have

$$D(X, d^n) = d^n(x_1, x_3) = d^n(x_2, x_3).$$

Another application of Theorem 2 gives a point x_4 such that

$$D(X, d^n) = a(X, d^n) = \frac{1}{3} (d^n(x_1, x_4) + d^n(x_2, x_4) + d^n(x_3, x_4)).$$

As before, this forces $D(X, d^n) = d^n(x_1, x_4) = d^n(x_2, x_4) = d^n(x_3, x_4)$.

Continuing in this fashion, we inductively obtain a sequence $x_1, x_2, ..., x_m, ...$ such that $d^n(x_i, x_j) = D(X, d^n)$. This sequence lies in a compact metric space, but has no convergent subsequence—which is impossible. Hence our supposition was false and $a(X, d^n) \neq D(X, d^n)$.

REMARK 2. It is obvious that m(X, f) always lies in the closed interval [-1, 1]. One cannot say more than this. For example, observe that if X is any compact connected Hausdorff space, then we can define $f: X \times X \to \mathbf{R}$ by f(x, y) = 1, for all x and y in X, yielding m(X, f) = 1.

A more interesting example is the following: let (X, d) be any compact connected metric space, and c any point of X. Define

$$f(x, y) = d(x, y) \cdot d(c, x) \cdot d(c, y), \text{ for } x, y \in X.$$

Then f is not identically zero, but m(X, f) = 0.

In his paper [13] Stadje also gives a formula for a(X, f):

(1)
$$a(X,f) = \sup_{\mu} \inf_{\nu} \int_{X} \int_{X} f(x,y) \, \mu(dx) \, \nu(dy),$$

where μ and ν run through $M^1(X)$, the set of all Borel probability measures on X. We shall prove this formula in §5.

The reader unfamiliar with Borel probability measures may choose to skip from here to §6.

The formula (1) is not very useful for calculating a(X, f) in general, for obvious reasons. However, Morris and Nickolas [11] were able to use it to calculate the magic numbers of spheres in \mathbb{R}^n and topological groups. They did this by first deducing the following as a consequence of (1).

THEOREM 3 [11]. Let X be a compact connected Hausdorff space and $f: X \times X \to \mathbf{R}$ a continuous symmetric function. If there exists a Borel probability measure μ_0 on X such that $\int_X f(x,y)\mu_0(dx)$ is independent of the choice of y in X, then $a(X,f) = \int_X f(x,y)\mu_0(dx)$ for any $y \in X$.

As noted by Morris and Nickolas [11, p. 463], an application of Theorem 3 to spheres Σ^n in

Euclidean space (with μ_0 being Lebesgue measure), shows that

$$m(\Sigma^n,d) = \frac{2^{n-1} \left[\Gamma\left(\frac{n+1}{2}\right) \right]^2}{\sqrt{\pi} \Gamma\left(\frac{2n+1}{2}\right)},$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$, the gamma function. (It is interesting to note that as $n \to \infty$, $m(\Sigma^n, d) \to 1/\sqrt{2}$.)

For topological groups, Morris and Nickolas [11] use Haar measure for μ_0 .

Rather than proving Theorem 3, we note that it is a corollary of the following result of Graham Elton:

THEOREM 4. If X is any compact connected Hausdorff space and $f: X \times X \to \mathbf{R}$ is a continuous symmetric function, then given any regular Borel probability measure μ on X, there is a point y in X such that $a(X, f) = \int_X f(x, y) \mu(dx)$.

This result of Elton clearly generalizes Stadje's theorem! It will be proved in §5 once we have shown that a(X, f) is unique. But first we point out another interesting application of Theorem 3.

In this example we calculate the magic number of any arc of a circle in the Euclidean plane. The method is that of Elton and depends upon Theorem 3. The "trick" is to find a measure μ_0 on the arc X such that $\int_X d(x, y) \mu_0(dx)$ is independent of the point y. [We thank Des Robbie for providing his calculations.]

EXAMPLE 5. Let X_{ϕ} be any arc of a circle with radius $\frac{1}{2}$ subtending an angle ϕ at the centre. Then $a(X_{\phi}, d) = \frac{4}{\phi + 4\cot\frac{\phi}{4}}$. So

$$m(X_{\phi},d) = \begin{cases} \frac{4}{\phi + 4\cot\frac{\phi}{4}}, & \text{if } \pi \leqslant \phi \leqslant 2\pi, \\ \frac{4}{\phi\sin\frac{\phi}{2} + 8\cos^2\frac{\phi}{4}}, & \text{if } 0 < \phi < \pi. \end{cases}$$

Proof. Let $P_{\theta} \in X_{\phi}$ be the point $(\frac{1}{2}\cos\theta, \frac{1}{2}\sin\theta)$, $0 \le \theta \le \phi$. Put $\mu_0 = \mu_1 + \mu_2 + \mu_3$, where μ_1 is 1-dimensional measure on the arc of the circle, μ_2 is a point measure giving weight m to P_0 , and μ_3 is a point measure giving weight m to P_{ϕ} . Then $\mu_0 / (\frac{\phi}{2} + 2m)$ is a probability measure on X_{ϕ} . (See Fig. 3.)

By Theorem 4 there is a point $P_{\alpha} \in X_{\phi}$ such that

$$a(X_{\phi},d) = \frac{\int_0^{\phi} d(P_{\alpha},P_{\theta}) \left(d\frac{1}{2}\theta\right) + m\left(\sin\frac{\alpha}{2} + \sin\frac{\phi - \alpha}{2}\right)}{\frac{\phi}{2} + 2m}.$$

So

$$a(X_{\phi},d) = \frac{2 + \left[m\sin\frac{\phi}{2} - \cos\frac{\phi}{2} - 1\right]\cos\frac{\alpha}{2} + \left[m - m\cos\frac{\phi}{2} - \sin\frac{\phi}{2}\right]\sin\frac{\alpha}{2}}{\frac{\phi}{2} + 2m}.$$

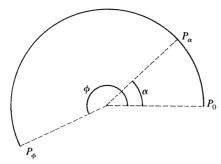


Fig. 3

Now, the term

$$\left(m\sin\frac{\phi}{2}-\cos\frac{\phi}{2}-1\right)\cos\frac{\alpha}{2}+\left(m-m\cos\frac{\phi}{2}-\sin\frac{\phi}{2}\right)\sin\frac{\alpha}{2}$$

will be independent of α if

$$\left(m\sin\frac{\phi}{2}-\cos\frac{\phi}{2}-1\right)=\left(m-m\cos\frac{\phi}{2}-\sin\frac{\phi}{2}\right)=0;$$

that is, when

$$m = \frac{\sin\frac{\phi}{2}}{1 - \cos\frac{\phi}{2}}$$

So we choose μ_2 , μ_3 to be the point measures which give weight

$$m = \frac{\sin\frac{\phi}{2}}{1 - \cos\frac{\phi}{2}}$$

to each endpoint. Then

$$a(X_{\phi},d) = \frac{4}{\phi + 4\cot\frac{\phi}{4}}.$$

When $0 < \phi < \pi$, the diameter of X_{ϕ} , $D(X_{\phi}, d) = \sin \frac{\phi}{2}$. So

$$m(X_{\phi},d) = \frac{4}{\phi \sin \frac{\phi}{2} + 8 \cos^2 \frac{\phi}{4}}, \quad \text{for } 0 < \phi < \pi.$$

When $\pi \leqslant \phi \leqslant 2\pi$, $D(X_{\phi}, d) = 1$, and therefore

$$m(X_{\phi},d) = \frac{4}{\phi + 4\cot\frac{\phi}{A}}.$$

(So if $\phi = 2\pi$, $m(X_{\phi}, d) = m(\Sigma^{1}, d) = 2/\pi$, which agrees with Example 4.)

While we have calculated the magic number of the circle, sphere, and arc of a circle, we do not know how to find the magic number of an ellipse.

OPEN QUESTION 2. What is the magic number of a general ellipse?

5. Uniqueness of a(X, f). In §3 we proved the existence of a(X, d) for any compact connected metric space (X, d) and in §4 indicated that this proof carried over to a(X, f). It is now time to prove the uniqueness of a(X, f). But first, it will be convenient to prove the existence part of Elton's Theorem 4.

THEOREM 4'. Let X and f be as in the statement of Theorem 4. Then there exists a real number a such that, given any regular Borel probability measure μ on X, there is a point $x \in X$ with $a = \int_X f(x, y) \mu(dy)$.

Proof. Equipped with the weak*-topology, $M^1(X)$, the set of all regular Borel probability measures on X, becomes a compact convex set. Its extreme points are the point masses corresponding to elements of X. So elements of \mathcal{F} , the set of all ordered n-tuples from X, may be regarded as convex combinations, with rational coefficients, of extreme points of $M^1(X)$. From the Krein-Milman theorem, we deduce that \mathcal{F} is dense in $M^1(X)$. (See Choquet [4, §25].)

Define $\Theta: X \times M^1(X) \to \mathbf{R}$ by $\Theta(x, \mu) = \int_X f(x, y) \mu(dy)$. (The number $\Theta(x, \mu)$ represents the distance from the point x to the probability measure μ .) It is routine to show that Θ is continuous. For $\mu \in M^1(X)$ let

$$a(\mu) = \min\{\Theta(x, \mu) : x \in X\}$$
 and $b(\mu) = \max\{\Theta(x, \mu) : x \in X\}$.

As in the previous existence proof, we have $\{\Theta(x, \mu): x \in X\} = [a(\mu), b(\mu)]$. The conclusion of this theorem may be stated as

$$(\exists a \in \mathbf{R}) (\forall \mu \in M^1(X)) a \in [a(\mu), b(\mu)].$$

As before, this can be the case only if $a(\mu) \leq b(\nu)$ for all $\mu, \nu \in M^1(X)$. Since $a_F \leq b_G$ for all F, $G \in \mathcal{F}$, the result follows from the density of \mathcal{F} in $M^1(X)$.

Proof of uniqueness of a(X, f). We continue the notation from the previous proof. That proof shows that uniqueness of a(X, f) is just the statement $\sup_{F \in \mathscr{F}} a_F = \inf_{G \in \mathscr{F}} b_G$, and that a(X, f) is their common value. From continuity of Θ , it follows that $a, b \colon M^1(X) \to \mathbb{R}$ are continuous functions. From the density of \mathscr{F} , it follows that

$$\sup_{F \in \mathscr{F}} a_F = \max_{\mu \in M^1(X)} a(\mu) \quad \text{and} \quad \inf_{G \in \mathscr{F}} b_G = \min_{\nu \in M^1(X)} b(\nu).$$

Now for $\mu, \nu \in M^1(X)$, let

$$A(\mu,\nu) = \int_{X} \int_{X} f(x,y) \mu(dy) \nu(dx).$$

For fixed μ , A is an affine continuous function of ν . By Bauer's maximum principle, [4, §25] it attains its minimum at an extreme point of $M^1(X)$. Thus

$$\min_{\nu} A(\mu, \nu) = \min_{\nu} \Theta(x, \mu) = a(\mu).$$

Similarly

$$\max_{\mu} A(\mu, \nu) = \max_{\nu} \Theta(y, \nu) = b(\nu),$$

and the uniqueness statement becomes

(2)
$$\max_{\mu} \min_{\nu} A(\mu, \nu) = \min_{\nu} \max_{\mu} A(\mu, \nu).$$

This is just Fan's version of the mini-max theorem [15, 6.3.8]. So uniqueness is proved. Also observe, now, that formula (2) yields formula (1).

REMARK 3. It is now clear that uniqueness of a(X, f) is essentially equivalent to the mini-max theorem, whereas existence of a(X, f) corresponds to the trivial mini-max inequality. This is why existence of a(X, f) is easier to prove than uniqueness.

Reexamining Examples 2 and 3 we observe that the technique used was to find measures μ and

 ν which maximized $a(\mu)$ and minimized $b(\nu)$. Since the resultant $a(\mu)$ and $b(\nu)$ are equal, their common value is a(X, d). This is the same as finding μ and ν for which

$$\Theta(x, \nu) \leq a(X, f) \leq \Theta(x, \mu)$$
, for all $x \in X$.

The above uniqueness proof shows that it is always possible to do this. (These μ and ν are called "optimal strategies".)

A question that arises naturally is: can we choose $\mu = \nu$? If so, we would have a measure μ for which $\Theta(x, \mu) = a(X, f)$ for all $x \in X$. This was essentially the idea used to calculate a(X, d) in Examples 4 and 5. The next result shows that it is not always possible to find such a probability measure. In particular, it is not possible in Examples 2 and 3. The idea of the following proof is due to David Wilson.

PROPOSITION 2. Let X be a compact connected subset of a rotund normed space. (This means that ||x + y|| < ||x|| + ||y|| unless x and y are linearly dependent. Observe that \mathbf{R}^n , for any n, is a rotund normed space.) Suppose that, for some probability measure μ on X, $\int_X ||x - y|| \mu(dy) = \Theta(x, \mu)$ is independent of x, and that X is not a line segment. Then no three points of X are collinear.

Proof. Suppose that, for some line L, $X \cap L$ contains at least three points. We show that X must be a line segment. Let a, b be two points in $X \cap L$ which are as far apart as possible, and let c be any other point in $X \cap L$. Then $c = \lambda a + (1 - \lambda)b$ for some $\lambda \in (0, 1)$. By rotundity,

$$||x - c|| < \lambda ||x - a|| + (1 - \lambda) ||x - b||,$$

for all $x \in X \setminus \{a, b\}$. Suppose μ is not concentrated on $\{a, b\}$. Integrating over X, we obtain

$$\Theta(c,\mu) < \lambda\Theta(a,\mu) + (1-\lambda)\Theta(b,\mu).$$

This contradicts our assumption that $\Theta(x, \mu)$ is independent of x. Thus μ is concentrated on $\{a, b\}$.

So

$$\Theta(x, \mu) = \mu\{a\} ||x - a|| + \mu\{b\} ||x - b||$$

for all $x \in X$. Since $\Theta(a, \mu) = \Theta(b, \mu)$, we must have $\mu\{a\} = \mu\{b\} = \frac{1}{2}$. Then, for any $x \in X$, we have

$$||x - a|| + ||x - b|| = 2\Theta(x, \mu) = 2\Theta(a, \mu) = ||a - b||.$$

By rotundity, x - a and x - b are linearly dependent. Thus X lies in the line segment

$$[a,b] = \{ \alpha a + (1-\alpha)b : 0 \le \alpha \le 1 \}.$$

Finally, connectedness forces X = [a, b].

6. The range of m(X, d). Often it is difficult to calculate a(X, f) precisely, and we must be content with some sort of estimate. In such cases, the following observation is surprisingly useful.

LEMMA 1 [17]. Suppose, for fixed α , $\beta \in \mathbb{R}$, that (X, f) has the following property: given $F \in \mathcal{F}$, there is a point $x \in X$ with $\alpha \leq \Theta(x, F) \leq \beta$. Then $\alpha \leq a(X, f) \leq \beta$.

Lemma 1 has already been used in some of the previous examples. Two further examples are given below.

EXAMPLE 6. Let (X, d) be the subspace of \mathbb{R}^2 given by

$$X = \left\{ (y, z) : y, z \in \mathbf{R}, \gamma^2 \leqslant y^2 + z^2 \leqslant \delta^2 \right\},\,$$

where γ and δ are positive real numbers. (See Fig. 4.)

Let $p = (0, \gamma)$. Then for all $x \in X$, $d(x, p) \le \gamma + \delta$. Hence, by Lemma 1 $a(X, d) \le \gamma + \delta$.

Example 7. Let $X = \left\{ \left(x, \sin \frac{1}{x} \right) : 0 < x \le \frac{1}{\pi} \right\} \cup \{ (0, y) : y \in [-1, 1] \}$. Then X is a com-

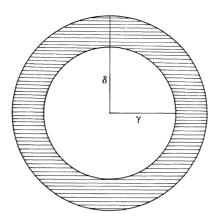


FIG. 4

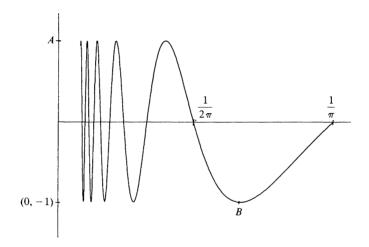


Fig. 5

pact connected subset of
$$\mathbb{R}^2$$
.
Let $x_1 = \left(\frac{1}{2\pi}, 0\right)$, $A = (0, 1)$, $B = \left(\frac{2}{3\pi}, -1\right)$. (See Fig. 5.)

For all $y \in X$,

$$\begin{split} d\big(x_1,y\big) & \leq d\big(x_1,A\big) \\ & = \left[1 + \frac{1}{4\pi^2}\right]^{1/2}. \end{split}$$

By Lemma 1

$$a(X, d) \leq \frac{1}{2\pi} [4\pi^2 + 1]^{1/2}.$$

Also

$$D(X,d) \ge d(A,B) = \left[4 + \frac{4}{9\pi^2}\right]^{1/2}$$
$$= \frac{2}{3\pi} [9\pi^2 + 1]^{1/2}.$$

Hence

$$m(X,d) = \frac{a(X,d)}{D(X,d)} \le \frac{\frac{1}{2\pi} [4\pi^2 + 1]^{1/2}}{\frac{2}{3\pi} [9\pi^2 + 1]^{1/2}}$$

\$\approx 0.5035.

So $0.500 \le m(X, d) \le 0.504$.

We shall now restrict our attention to the metric case, and investigate the range of values possible for m(X, d), when various restrictions are placed on (X, d).

Gross [9] showed that m(X, d) always lies in the half-open interval $\left[\frac{1}{2}, 1\right)$. This is a special case of Proposition 1. If m(X, d) is close to 1, this indicates that "most" points of X are far apart from one another. On the other hand, it is possible to have $m(X, d) = \frac{1}{2}$ for a space in which "many" points are far apart from one another. As an example of this phenomenon, observe that the magic number for any circle is $2/\pi$ (Example 4), whereas the magic number for a circle together with any of its diameters is $\frac{1}{2}$ (use the method of Example 2). More generally, the following is true.

PROPOSITION 3. Let (X, d) be any compact connected metric space, and choose m with $\frac{1}{2} \le m < m(X, d)$. Then there exists a wedge $X \lor I$ (that is, a space obtained from X by gluing on a line segment I) such that the magic number of $X \lor I$ equals m.

Proposition 3 was first proved in [17], but a simpler proof can be found in [6].

NOTATION. For each metric space (S, d), we define g(S) to be the supremum of the numbers m(X, d) as X ranges over all compact connected subsets of S.

If the supremum is attained, then, by Proposition 1, it is strictly less than one. This is the case, for example, for any finite-dimensional normed space [17, Theorem 5]. (If E is an infinite dimensional normed space, then g(E) = 1, by Dvoretzky's theorem [7] and [17, Theorem 9]. In this case, by Proposition 1, the supremum cannot be attained.)

OPEN QUESTION 3. Let \mathbb{R}^n denote Euclidean *n*-space. Then what is $g(\mathbb{R}^n)$? Indeed what is $g(\mathbb{R}^2)$?

Gross [9] stated that "computations seem to indicate that the bound (that is, $g(\mathbf{R}^2)$) is not much greater than 2/3". We suggest that $g(\mathbf{R}^2)$ may equal the magic number of the Reuleaux triangle. (The Reuleaux triangle consists of the vertices of an equilateral triangle together with three arcs of circles, each circle having centre at one of the vertices and endpoints, the other two vertices. See Fig. 6.)

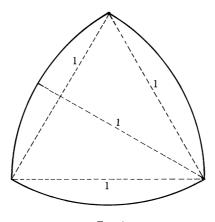


Fig. 6

The magic number for the Reuleaux triangle is not known exactly, but computer approximation shows that it is about 0.668.

OPEN QUESTION 4. What is the magic number of the Reuleaux triangle?

In 1982 Graham Elton (private communication, see also [5]) proved that $g(\mathbf{R}^2) \leq 0.775$. More recently Peter Nickolas [12] has proved that $g(\mathbf{R}^2) \leq 4/\pi\sqrt{3} \approx 0.735$. From this it follows that $0.667 < g(\mathbf{R}^2) < 0.736$. Nickolas does this by proving that if (X, d) is any compact connected metric subspace of a disk of diameter one in \mathbf{R}^2 , then $a(X, d) \leq 2/\pi$, which is, of course, the magic number of the circle of diameter one. Then, using the fact that every subset of \mathbf{R}^2 of diameter not greater than one is a subset of a disk of diameter $2/\sqrt{3}$, one obtains that for any compact connected metric subspace (Y, d) of \mathbf{R}^2

$$m(Y,d) \leqslant \frac{2}{\pi} \cdot \frac{2}{\sqrt{3}} = \frac{4}{\pi\sqrt{3}}.$$

Using a similar technique, but with the disk replaced by a hexagon, the third author has recently improved this estimate to $g(\mathbf{R}^2) \leq \frac{2+\sqrt{3}}{3\sqrt{3}} \simeq 0.718$. This uses the fact that any set in \mathbf{R}^2 of diameter one is contained in a regular hexagon of side length $1/\sqrt{3}$.

While we do not know the exact value of $g(\mathbb{R}^2)$, the analogous problem for *convex* subsets of \mathbb{R}^2 (indeed \mathbb{R}^n) has been solved. To attack this problem, it is appropriate to consider convex sets in general normed linear spaces.

NOTATION. Let X be a compact convex subset of some normed space. We denote by r(X) the minimum of the radii of those closed balls, with centres in X, which contain X. A simple compactness argument shows that this minimum exists.

It is essential in this definition to consider only balls with centres in X. A result of Garkavi [8] asserts that if E is a normed space of dimension at least three, which does not admit an inner-product structure, then E contains a compact convex set X, which is contained in a ball of radius strictly less then r(X). See Borwein and Keener [3].

The following elegant result is due to Esther and George Szekeres (private communication).

THEOREM 5. Let X be a compact convex subset of some normed space, with d being the metric given by the norm. Then a(X, d) = r(X).

Proof (Szekeres). Let x_0 be the centre of a ball, of radius r = r(X), which contains X. Then $d(x, x_0) \le r$ for all $x \in X$. Choose $x_1, \ldots, x_n \in X$ and define $\Theta: X \to \mathbf{R}$ by

$$\Theta(x) = \frac{1}{n} \sum_{i=1}^{n} d(x, x_i).$$

We must show that Θ takes the value r somewhere on X. Clearly $\Theta(x_0) \le r$. Since X is connected, it suffices to show that $\Theta(x) \ge r$ for some $x \in X$.

We let $c = \frac{1}{n} \sum_{i=1}^{n} x_i \in X$, and choose $y \in X$ so that $d(c, y) \ge r$. At least one such y exists, by definition of r(X). Then

$$r \le ||c - y|| \le \frac{1}{n} \sum_{i=1}^{n} ||x_i - y|| = \Theta(y),$$

as required.

We remark that this proof does not use the mini-max technique so common in previous examples.

We can now calculate the convex analogue of $g(\mathbf{R}^n)$.

NOTATION. For any normed space E, let s(E) be the supremum of the numbers m(X, d) as X

ranges over all compact convex subsets of E.

THEOREM 6 [14]. For each positive integer n, $s(\mathbf{R}^n) = \sqrt{n/(2n+2)}$.

Proof. A classical result of Jung [10] states that if X is a compact convex subset of \mathbb{R}^n , with unit diameter, then $r(X) \leq \sqrt{n/(2n+2)}$. Blumenthal and Wahlin [2] gave a simple proof of this, and also showed that the upper bound is attained when X is the regular n-simplex. The result now follows from Theorem 5.

For recent work relating to Jung's theorem, we refer the reader to [1], and the references therein.

REMARK 4. The proof of Theorem 6 first given by Strantzen [14] is direct, and independent of Theorem 5. He also calculated the magic numbers for the k-skeletons of a regular n-simplex, $1 \le k \le n$. (The k-skeleton of an n-simplex X is the union of all k-simplices whose vertices are already vertices of X.) Stadje [13] proved that $s(\mathbf{R}^2) \le \frac{1}{2}\sqrt{5-2\sqrt{3}}$, a weaker result than Theorem 6. He claimed that the same bound was valid for $s(\mathbf{R}^n)$, but, as shown by Strantzen [14], this is false for $n \ge 4$.

REMARK 5. Note that as n tends to infinity, $s(\mathbf{R}^n)$ tends to $1/\sqrt{2}$.

REMARK 6. It is shown in [17] that the maximum of s(E) as E ranges over all n-dimensional normed spaces, is n/(n+1).

7. Numerical geometry...not numerical topology. In §6 we remarked that the magic number provides some information on how "spread out" a space is. The following surprising result shows that the magic number depends heavily on the metric, not simply on the topology of the space. So it is appropriate to call this subject numerical geometry, rather than numerical topology.

THEOREM 7 [6]. Let X be a compact connected metrizable space. Then for each real $m \in [\frac{1}{2}, 1)$, there is a compatible metric d on X such that m(X, d) = m.

Outline Proof. Choose a compatible metric ρ on X so that $D(X, \rho) = 1$ and then choose $a, b, c \in X$ with $\rho(a, b) = 1$ and $\rho(a, c) = \frac{1}{2}$. Then define d by

$$d(x, y) = \min \left\{ \rho(x, y), \min \left\{ \frac{1}{2}, \rho(c, x) \right\} + \min \left\{ \frac{1}{2}, \rho(c, y) \right\} \right\}.$$

It can be verified that d is a metric, equivalent to ρ , and that $m(X, d) = \frac{1}{2}$. For the case $m > \frac{1}{2}$, consider the family of metrics $d_{\lambda}(\lambda \ge 0)$ defined by

$$d_{\lambda} = (\lambda + 1) d/(\lambda d + 1).$$

Then $D(X, d_{\lambda}) = 1$ for all λ , and if $x \neq y$, then $d_{\lambda}(x, y) \to 1$ as $\lambda \to \infty$. Some straightforward but tedious analysis then shows that $m(X, d_{\lambda})$ is a continuous function of λ , and that $m(X, d_{\lambda}) \to 1$ as $\lambda \to \infty$. Since $m(X, d_{0}) = \frac{1}{2}$, it follows from the Intermediate Value Theorem that $m(X, d_{\lambda}) = m$ for some λ .

For compact connected Hausdorff spaces, the following analogue of Theorem 7 holds. We give a new proof.

THEOREM 8 [6]. Let X be a compact connected Hausdorff space. Then for each $m \in [-1,1]$ there is a continuous symmetric function $f: X \times X \to \mathbf{R}$ such that m(X, f) = m.

Proof. Firstly let $m \in [0,1]$. Choose distinct points a, b, c in X and put

$$S = (X \times \{c\}) \cup (\{c\} \times X) \cup \{(a,b)\} \cup \{(b,a)\}.$$

So S is a closed subspace of the compact Hausdorff space $X \times X$. Define a continuous function $g: S \to [0,1]$ as follows: $g(x,c) = g(c,x) = \sqrt{m}$, for all $x \in X$, and g(a,b) = g(b,a) = 1. By

Tietze's extension theorem, there exists a continuous function θ : $X \times X \to [0,1]$ such that $g(s) = \theta(s)$, for all $s \in S$. Now define f: $X \times X \to [0,1]$ by $f(x,y) = \theta(x,y)\theta(y,x)$, for all $(x,y) \in X \times X$. Then f is a continuous symmetric function with the property that f(x,c) = m, for all $x \in X$. From this it immediately follows that a(X,f) = m. As D(X,f) = 1, we have m(X,f) = m, as required.

If $m \in [-1,0]$, then we find, as above, a function f such that m(X,f) = -m. Then putting f' = -f, we have m(X,f') = m.

8. The squared distance function. As indicated in Remark 1 the calculation of $a(X, d^2)$ for subspaces (X, d) of Euclidean space, \mathbb{R}^n , is much simpler than that of a(X, d). We shall only touch on the topic here as full details appear in the paper [16] of David Wilson. The key theorem is:

THEOREM 9. Let X be a compact connected subset of \mathbb{R}^n . Let B_1 be a closed ball and B_2 an open ball such that $X \subset B_1 \setminus B_2$ and the centre of each ball lies in the closed convex hull of the intersection of X with the boundary of the other. Further, let B_1 have centre α and radius R and let B_2 have centre β and radius R. Then $a(X, d^2) = R^2 + r^2 - |\alpha - \beta|^2$.

If we consider compact *convex* subsets X of \mathbb{R}^n , then by the hypotheses of Theorem 9, r must be zero and α must equal β . Therefore $a(X, d^2) = R^2$. Using a modification of the Szekeres proof (Theorem 5) that the magic number of a compact convex set is the circumradius of the set, Wilson shows that R is the circumradius of X.

Open Question 2 asks for the magic number of an ellipse. The analogous problem for the squared distance function is easy using Theorem 9.

EXAMPLE 8 [16]. Let X be the ellipse with semi-major axis of length R and semi-minor axis of length r. Clearly $\alpha = \beta$, and so $a(X, d^2) = R^2 + r^2$. (See Fig. 7.)

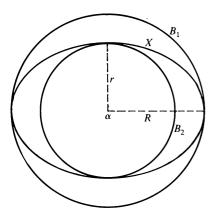


Fig. 7

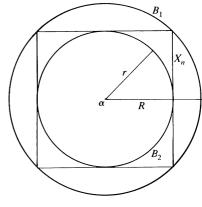


Fig. 8

The regular n-gon is also easily handled:

EXAMPLE 9 [16]. Let X_n be the regular *n*-sided polygon with vertices on the circle of radius $\frac{1}{2}$ and centre the origin. (See Fig. 8.)

Again
$$\alpha = \beta$$
. Now $R^2 = \frac{1}{4}$ and $r^2 = \frac{1}{8} \left(1 + \cos \frac{2\pi}{n} \right)$. Therefore $a(X, d^2) = \frac{1}{8} \left(3 + \cos \frac{2\pi}{n} \right)$.

Open Question 4 asks for the magic number of the Reuleaux triangle. Wilson [16] shows that

$$m(X, d^2) = \frac{5 - 2\sqrt{3}}{3}$$
.

He also observes that by deleting one of the sides of the Reuleaux triangle, one obtains a space Y with

$$m(Y, d^2) = \frac{1}{2} \left(3 - \sqrt{\frac{\pi}{3}}\right)$$

which is greater than the value for the Reuleaux triangle itself.

OPEN QUESTION 5. What is the value of $g_2(\mathbf{R}^n)$, the supremum of the numbers $m(X, d^2)$ as X ranges over all compact connected subsets of \mathbf{R}^n ?

References

- 1. D. Amir and C. Franchetti, The radius ratio and convexity properties in normed linear spaces, Trans. Amer. Math. Soc., 282 (1984) 275–292.
- 2. L. M. Blumenthal and G. E. Wahlin, On the spherical surface of smallest radius enclosing a bounded subset of *n*-dimensional Euclidean space, Bull. Amer. Math. Soc., 47 (1941) 771–777.
- 3. J. Borwein and L. Keener, The Hausdorff metric and Čebyšev centres, J. Approx. Theory, 28 (1980) 366-376.
 - 4. G. Choquet, Lectures on Analysis, vol. II, Benjamin, Reading, Massachusetts, 1969.
- 5. J. Cleary, Numerical Geometry, Honours project; La Trobe University, 1982 (La Trobe University Pure Mathematics Research Paper 83-2).
- 6. J. Cleary and S. A. Morris, Numerical geometry...not numerical topology, Math. Chronicle, 13 (1984) 47-58.
- 7. T. Figiel, A short proof of Dvoretzky's theorem on almost spherical sections of convex bodies, Compositio Math., 33 (1976) 297–301.
- 8. A. L. Garkavi, On the Chebyshev centre and convex hull of a set, Uspehi Math. Nauk, 19 (1964) no. 6. (120) 139–145.
- 9. O. Gross, The rendezvous value of a metric space, in Advances in game theory, Ann. of Math. Studies no. 52, Princeton, 1964, 49–53.
- 10. H. Jung, Ueber die kleinste Kugel, die eine räumliche Figur einschliesst, J. Reine Angew. Math., 123 (1901) 241–257.
- 11. S. A. Morris and P. Nickolas, On the average distance property of compact connected spaces, Arch. Math., 40 (1983) 459–463.
 - 12. P. Nickolas, On the average distance property for subsets of the plane, to appear.
 - 13. W. Stadje, A property of compact, connected spaces, Arch. Math., 36 (1981) 275-280.
 - 14. J. Strantzen, An average distance result in Euclidean n-space, Bull. Austral. Math. Soc., 26 (1982) 321–330.
 - 15. J. Stoer and C. Witzgall, Convexity and Optimization in Finite Dimensions I, Springer-Verlag, Berlin, 1970.
- 16. D. J. Wilson, A game with squared distance as payoff, Melbourne University Mathematics Research Report 1982, no. 16.
 - 17. D. Yost, Average distances in compact connected spaces, Bull. Austral. Math. Soc., 26 (1982) 331-342.

165.

MISCELLANEA

If you see someone walking about in the vicinity of a campus who looks as if he's either very backward or very brilliant, then if he's not backward he's a mathematician. There is not another group of people, I'll wager, with eccentricities so pronounced and pure, with personalities so undiluted by the attempt to conform.

— Rebecca Goldstein, The Mind-Body Problem: a Novel, Andre Deutsch, 1985.

ANSWER TO PHOTO ON PAGE 259

Alexandra Bellow.