71-20,936

FRIEDMAN, Susan Loeserman, 1934-
GENERALIZED TRANSFINITE DIAMETERS AND CHEBYSHEV
CONSTANTS.

The City University of New York, Ph.D., 1971
Mathematics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

# GENERALIZED TRANSFINITE DIAMETERS AND CHEBYSHEV CONSTANTS

by

Susan Loeserman Friedman

A dissertation submitted to the Graduate
faculty in Mathematics in partial fulfillment
of the requirements for the degree of Doctor
of Philosophy, The City University of New York.

1971

This manuscript has been read and accepted for the Graduate Faculty in Mathematics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

JAN. 27, 1971
date

Theodore J. Rivlin
Chairman of Examining Committee

1/27/71
date

Executive Officer

Professor Alan J. Hoffman

Professor Linda Keen

Professor Richard Sacksteder

Supervisory Committee

The City University of New York

# TABLE OF CONTENTS

# ACKNOWLEDGEMENTS

Every Ph.D. degree is equivalent to a finite sum of weighted terms, the number of which is directly proportional to the candidate's age when he finally achieves the degree. The set of terms is never empty, for the ever-present term and the one with the greatest weight is one's dissertation advisor. It has been my good fortune to have Professor Theodore J. Rivlin for my advisor. I wish to express my gratitude to him for his guidance, encouragement, and friendship over the years. Many, many times he has cheered me on.

Another term of considerable magnitude in this sum, as well as foremost in my life, is my husband, Stanley D. Friedman, without whose love, patience, reassurance, and belief that women, too, have a right to seek fulfillment, nought would have been accomplished. And of course my thanks go to my children, - Deborah, Daniel, and Michael, who, in more or less equally weighted fashion, suffered through a long summer of silence while mother studied for her qualifying examination.

I am very much indebted to my housekeeper, Mrs. Mildred C. Grant, who over the past six years has kept things at home on an even keel - a substitute mother in every sense of the word so that my children, I feel, were not damaged by my mental and physical absence.

Also present and of goodly weight in the sum are my parents, Arthur and Ruth Loeserman, and my father-in-law Joseph N. Friedman (my high school mathematics teacher). I have always been mindful of their love and sustained by their belief in my being able to persevere.

I also wish to express my appreciation to Professor Leo Zippin for his aid in obtaining financial help for several years during this

undertaking, and to Mrs. Ione Hutson for her painstaking care in typing this manuscript.

And perhaps in the mystical summation sign is the candidate herself, fortunate in being able to combine all the terms into a meaningful whole...

# INTRODUCTION

The transfinite diameter and the Chebyshev constant are two non-negative set functions defined for compact sets in the complex plane (cf. Fekete [6]). They are limits of sequences, and for each n in the sequence we have associated with these concepts, sets of n points termed Fekete points ( $n \geq 2$ ), and Chebyshev points ( $n \geq 1$ ), respectively. Monic polynomials of degree n , whose zeros are n-Fekete points ( n-Chebyshev points) are termed Fekete polynomials (Chebyshev polynomials).

Chapter 1 deals with the classical theory and provides the reader with the basic theorems of the subject. Among these is the Convex Hull Theorem (Theorem 1.2.3), which is proved using Fejér's Principle. It provides information about the location of Chebyshev points. Another important theorem is Fekete's Theorem (Theorem 1.3.1) which establishes the equality of the transfinite diameter and the Chebyshev constant for all compact sets in the complex plane. In this chapter the reader is provided with examples to illustrate the concepts defined. The only original material here (Theorem 1.5.4) deals with the related concept of "restricted" Chebyshev points and "restricted" Chebyshev polynomial.

Further investigations were made in the complex plane using more general metrics to define the transfinite diameter and the Chebyshev constant, see Frostman [7] and Tsuji [32]. Pólya and Szegö [25] used different averaging processes - notably $\lambda$-th power means - to find the transfinite diameter and the Chebyshev constant for certain compact sets on the real line, in the complex plane, and in three-dimensional

Euclidean space. Their work included the classical case, since the zero-th power means (geometric average) was the original average used by Fekete. Hille [16] generalized further by using averages based on a set of Postulates formulated by Kolmogorov [18] and Nagumo [32]. A discussion of averaging processes and the theorems dealing with them, including the postulates used by Hille is given in Chapter 2.

Hille's averages are generated by functions through the correspondence given in equation (2.1.5). It is through these generating functions that we generalize the concepts of transfinite diameter and Chebyshev constant. This approach is different, but no substantially new material results, as it is equivalent to working through the averages themselves, as Hille did. However, one can use this approach to generalize the classical results of Chapter 1. In this new setting we examine the location of generalized Fekete points and generalized Chebyshev points in arbitrary metric spaces.

The two important theorems of Chapter 1 no longer hold. In Theorem 2.3.3 we extend the work of Shisha [27] with a theorem dealing with Fejér's Principle in inner-product spaces. Thus we prove Theorem 2.3.4, a Convex Hull theorem for Hilbert space. Fekete's equality in Theorem 1.3.1 does not hold in an arbitrary metric space and must be replaced by an inequality. This known result is proved here via generating functions for averages. The concluding section of Chapter 2 deals with some theorems about the relationship between generalized restricted and generalized unrestricted Chebyshev points. Although these results follow almost directly from the definitions, they have not been published before.

Chapter 3 contains the heart of this dissertation and the bulk of the original material contained therein. Restricting ourselves to $\lambda$-th power averages, $\lambda \geq 1$, it is here that we develop methods to characterize Chebyshev points for all compact sets in real Euclidean n-space.

In $R_1$, the real line, we locate Fekete points and Chebyshev points for any compact set. We examine the related question of their unicity as well as discuss the generalized lemniscatic regions arising from their associated Fekete functions and Chebyshev functions. Finally we discuss the cases of equality (for $\lambda = 1$) and strict inequality (for $\lambda > 1$) between the transfinite diameter and the Chebyshev constant.

In $R_1$ it is found the Chebyshev constant for any compact set E for $\lambda$-th power averages, $\lambda \geq 1$, is equal to $\frac{1}{2}(b - a)$ where $b = \max_{x \in E} x$ and $a = \min_{x \in E} x$. This value could be interpreted as the radius of a two-point spanning sphere of the set E. This observation led us to employ the known concept of spanning circle (see Yaglom [33]) in two-dimensional Euclidean space. Together with a solution of a generalized Steiner problem by means of a characterization theorem in approximation theory, we are able to prove the main theorem in Section 3.2 (Theorem 3.2.5) showing that the Chebyshev constant for any compact set for $\lambda$-th power averages, $\lambda \geq 1$ is equal to R, where R is the radius of the spanning circle of E. Topics such as the location of Chebyshev points, the question of their unicity, and the comparison between the transfinite diameter and the Chebyshev constant are then discussed.

The concluding section of Chapter 3 generalizes all the material used to prove the main theorem of Section 3.2 to n-dimensional real Euclidean space.

Chapter 4 deals with related topics found in other material in a different context. Section 4.1 shows that we can find the Chebyshev constant for $\lambda$-th power averages, $\lambda \geq 1$ for all centered subsets of metric spaces and for all subsets of centered spaces. Hence we examine various metric spaces for centeredness and quote an important theorem of Kolmogorov and Tihimirov [19]. The fact that the unit ball is centered leads us to an upper bound of one for the Chebyshev constant of that set for all averaging processes satisfying Hille's postulates. Thus, using Hille's results for the transfinite diameters of unit balls in certain metric spaces, we can establish strict inequality between the two set functions.

Section 4.2 deals with Chebyshev centers. As defined in Singer [28], a Chebyshev center is merely what we called a Chebyshev point of degree one. We quote the results he cites dealing with existence and uniqueness of Chebyshev centers, and we establish two theorems dealing with the relationship between centers and Chebyshev centers.

The concluding section, Section 4.3 deals with another concept quoted in Singer called "closest points." We show how this concept relates to Fejér's Principle. Theorems found in Phelps [24] give certain conditions which imply Fejér's Principle. Hence this gives us additional information about spaces in which the Convex Hull theorem holds true.

# CHAPTER 1

## The Classical Theory

The classical setting for the topics discussed in this paper is the complex plane, $\mathbb{C}$ . In 1923, while investigating an algebraic problem, M. Fekete [6] defined a set function which he called the transfinite diameter. He proceeded as follows:

### 1.1 The Transfinite Diameter

Let E be a compact set in $\mathbb{C}$ containing infinitely many points. Let $z_1, z_2, \ldots, z_n$ be points of E, $n \geq 2$ and consider

$$V(z_1, z_2, \ldots, z_n) = \prod_{1 \leq i < j \leq n} |z_i - z_j| \qquad (1.1.1)$$

If the $z_i$'s are distinct, the product is not zero. It is a continuous function of n variables and on the compact set E will attain its maximum. Set

$$V_n(E) = \max_{z_1, z_2, \ldots, z_n \in E} V(z_1, z_2, \ldots, z_n) \ . \qquad (1.1.2)$$

Let

$$d_n(E) = (V_n(E))^{\frac{1}{\binom{n}{2}}} \qquad (1.1.3)$$

**Theorem 1.1.1**

The sequence $\{d_n(E)\}$ is monotone decreasing.

**Proof:** See Goluzin [10].

Since $\{d_n(E)\}$ is a bounded (by zero), monotone, decreasing sequence it converges and we have

**Definition 1.1.1**

$\delta(E) = \lim_{n \to \infty} d_n(E)$ is called the transfinite diameter of the set E .

We note that $d_2(E) = \max_{z_1, z_2 \in E} |z_1 - z_2| = d(E)$ , the topological diameter of E . Thus we have

$$\delta(E) \leq d(E) \ . \tag{1.1.4}$$

If E contains finitely many points, say k , then for $n > k$ $V(z_1, z_2, \ldots, z_n) = 0$ and hence $\delta(E) = 0$ .

The transfinite diameter of a set has the properties of monotony and continuity.

Theorem 1.1.2

a) If $E_1 \subset E_2$ , then $\delta(E_1) \leq \delta(E_2)$ . (1.1.5)

b) Let $E_\epsilon$ be the set of points of $\mathbb{C}$ having a distance from E not exceeding $\epsilon$ . Then $\lim_{\epsilon \to 0} \delta(E_\epsilon) = \delta(E)$ . (1.1.6)

In addition we can establish the following effect on the transfinite diameter caused by certain types of mappings.

Theorem 1.1.3

Let E be a compact set in $\mathbb{C}$ .

a) Let $F = \{az+b \,|\, z \in E\}$ . Then $\delta(F) = |a|\delta(E)$ . (1.1.7)

b) If $q(z) = z^k + a_1 z^{k-1} + \ldots + a_k$ and $F = \{z \,|\, q(z) \in E\}$ , then $\delta(F) = [\delta(E)]^{\frac{1}{k}}$ (1.1.8)

See Hille [14] for proofs of above theorems.

Definition 1.1.2

For every n , we will associate with the set E a polynomial of degree n whose zeros will be $\xi_{n,1}, \xi_{n,2}, \ldots, \xi_{n,n}$ such that $\max_{z_1, z_2, \ldots, z_n \in E} V(z_1, z_2, \ldots, z_n) = V(\xi_{n,1}, \xi_{n,2}, \ldots, \xi_{n,n})$ . This polynomial will be called a <u>Fekete polynomial of degree n for E</u> and will be written $F_n(z;E) = \prod_{i=1}^{n} (z-\xi_{n,i})$ .

<u>Definition 1.1.3</u>

The set of points $\xi_{n,1}, \xi_{n,2}, \ldots, \xi_{n,n}$ will be termed <u>Fekete points of order n for the set E</u> .

Of course, if the Fekete points of order $n$ for $E$ are unique, then the Fekete polynomial of degree $n$ for $E$ will be unique.

<u>Definition 1.1.4</u>

Let $P(z)$ be a polynomial and consider the set $\{z \mid |P(z)| = R\}$ where $R$ is a non-negative constant. This set is called a <u>lemniscate</u>. The set $\{z \mid |P(z)| \le R\}$ is called a <u>lemniscatic region</u>.

With each Fekete polynomial $F_n(z;E)$ of degree $n$ , we can define a lemniscatic region $\mathscr{L}_{F_n}$ in the following manner.

Let
$$K_n = \|F_n(z;E)\|_E = \max_{z\in E} |F_n(z;E)| . \tag{1.1.9}$$

Then by $\mathscr{L}_{F_n}$ we shall mean

$$\{z \mid |F_n(z;E)| \le K_n\} \tag{1.1.10}$$

and by $\partial\mathscr{L}_{F_n}$ we shall mean the lemniscate

$$\{z \mid |F_n(z;E)| = K_n\} \tag{1.1.11}$$

Theorem 1.1.4

a) $E \subset \mathcal{L}_{F_n}$ for every $n$ .

b) $\partial\mathcal{L}_{F_n} \cap E \neq \emptyset$ .

c) $\lim_{n\to\infty} K_n^{\frac{1}{n}} = \delta(E)$ .

Proof: See Hille [14].

Definition 1.1.5

Consider the extended complex plane. Since $E$ is compact, its complement $E^c$ is open and has at most countably many connected components. One of these components contains the point at $\infty$ . We shall call that component $E^c_\infty$ . The common boundary of $E^c_\infty$ and $E$ , $\overline{E^c_\infty} \cap E$ , we call the outer boundary of $E$ .

To prove that the Fekete points of $E$ lie on the outer boundary of $E$ , we consider the basic polynomials:

$$F_{n,j} = \frac{(z-\xi_{n,1})(z-\xi_{n,2})\cdots(z-\xi_{n,j-1})(z-\xi_{n,j+1})\cdots(z-\xi_{n,n})}{(\xi_{n,j}-\xi_{n,1})(\xi_{n,j}-\xi_{n,2})\cdots(\xi_{n,j}-\xi_{n,j-1})(\xi_{n,j}-\xi_{n,j+1})\cdots(\xi_{n,j}-\xi_{n,n})}$$

for $j = 1,2,\ldots,n$ (1.1.12)

where $\xi_{n,1},\xi_{n,2},\ldots,\xi_{n,n}$ are Fekete points of order $n$ for $E$ . (Polynomials of this form are used for polynomial interpolation at $\xi_{n,1},\xi_{n,2},\ldots,\xi_{n,n}$ , but we are not interested in that problem here.) We note

$$F_{n,j}(\xi_{n,i}) = 0 \quad \text{for} \quad i \neq j$$

$$F_{n,j}(\xi_{n,j}) = 1 \qquad (1.1.13)$$

**Theorem 1.1.5**

$$\|F_{n,j}\|_E = \max_{z\in E} |F_{n,j}(z)| = 1$$

**Proof:** See Hille [14].

As a consequence of the maximum modulus principle and with the help of Theorem 1.1.5 we have the result:

**Theorem 1.1.6**

The Fekete points $\{\xi_{n,j}\}$ lie on the outer boundary of $E$ .

**Proof:** See Hille [14].

Theorem 1.1.6 shows that the transfinite diameter of a set $E$ is equal to the transfinite diameter of the set $\overline{E^c_\infty} \cap E$ , the outer boundary of $E$ . Intuitively, the removal of open spheres or the introduction of "holes" in a set $E$ does not alter the transfinite diameter of $E$ .

## 1.2 The Chebyshev Constant

Besides Fekete polynomials, we consider another class of polynomials associated with a compact set $E$ in $\mathbb{C}$ , called Chebyshev polynomials. Let $\mathcal{P}_n$ be the class of monic polynomials of degree $n$ , i.e., $P_n(z) \in \mathcal{P}_n$ if and only if $P_n(z) = z^n + a_1 z^{n-1} + \ldots + a_n$ where $a_i$, $i = 1,2,\ldots,n$ are complex numbers. Since $E$ is compact there is a point $z$ in $E$ such that $|P_n(z)|$ is maximal. The set of maxima for all monic polynomials is bounded below and the infimum is reached.

**Definition 1.2.1**

The polynomial in $\mathcal{P}_n$ whose maximal absolute value is minimal over all polynomials in $\mathcal{P}_n$ is called the *Chebyshev polynomial of degree $n$ for $E$* and we denote it by $T_n(z)$ or $T_n(z;E)$ .

Theorem 1.2.1

There exists a monic polynomial $T_n(z) \in \mathcal{P}_n$ such that

$$m_n = \max_{z \in E} |T_n(z)| = \min_{P_n \in \mathcal{P}_n} \max_{z \in E} |P_n(z)|$$

Proof: See Tsuji [32].

If the cardinality of $E$ is $k$ where $k < n$, then $m_n(E) = 0$, and any polynomial of degree $n$ vanishing on $k$ points has the property of the Chebyshev polynomial. On the other hand, if $E$ consists of $n$ or more points, then one can show:

Theorem 1.2.2

$T_n(z;E)$ is unique.

Proof: See Tsuji [32].

Since we often define the norm of a polynomial on $E$ by $\|P_n(z)\|_E = \max_{z \in E} |P_n(z)|$, we see that the Chebyshev polynomial of degree $n$ for a set $E$ is the monic polynomial of minimum norm for the set. Thus

$$m_n(E) = \max_{z \in E} |T_n(z)| = \|T_n(z)\|_E = \|T_n(z;E)\| \qquad (1.2.1)$$

Let

$$\tau_n(E) = [m_n(E)]^{\frac{1}{n}} \qquad (1.2.2)$$

and consider $\lim_{n\to\infty} \tau_n(E)$. This limit exists (see Goluzin [10]) and we have:

Definition 1.2.2

$\chi(E) = \lim_{n\to\infty} \tau_n(E)$ is called the Chebyshev constant for $E$.

We note that $m_n(E) = \|T_n(z;E)\| \le \|(z-z_0)^n\|$ for every $z_0$ in $E$.

Thus $\tau_n(E) \le (\max_{z\in E} |z-z_0|^n)^{\frac{1}{n}} = \max_{z\in E} |z-z_0| \le d(E)$ , the topological diameter of E . Hence

$$\chi(E) \le d(E) \tag{1.2.3}$$

Let $z_1, z_2, \ldots, z_n$ be the zeros of $P_n(z)$ . Exhibiting the polynomial in its product form we can write $P_n(z) = \prod_{i=1}^{n} (z-z_i)$ and $|P_n(z)| = |\prod_{i=1}^{n} (z-z_i)| = \prod_{i=1}^{n} |z-z_i|$ . Then

$$\tau_n(E) = [m_n(E)]^{\frac{1}{n}} = \max_{z\in E} |T_n(z;E)|^{\frac{1}{n}} = \min_{P_n\in\mathcal{P}_n} \max_{z\in E} |P_n(z)|^{\frac{1}{n}} \quad or$$

$$\tau_n(E) = \min_{\substack{z_i\in\mathbb{C} \\ i=1,2,\ldots,n}} \max_{z\in E} |\prod_{i=1}^{n} (z-z_i)|^{\frac{1}{n}} = \min_{\substack{z_i\in\mathbb{C} \\ i=1,2,\ldots,n}} \max_{z\in E} \prod_{i=1}^{n} |z-z_i|^{\frac{1}{n}} \tag{1.2.4}$$

Thus one can think of $\tau_n(E)$ as either 1) the n-th root of the norm of the Chebyshev polynomial of degree n for E , or as 2) the minimal value of the maximum geometric mean of the product of the distances from a variable point z in E to n points $z_1, z_2, \ldots, z_n$ in the complex plane.

Although both representations are equivalent, it is the latter that we shall deal with more often in the future.

<u>Definition 1.2.3</u>

Let $z_1^*, z_2^*, \ldots, z_n^*$ be such that $m_n(E) = \max_{z\in E} \prod_{i=1}^{n} |z-z_i^*|$ . We shall call $z_1^*, z_2^*, \ldots, z_n^*$ <u>Chebyshev points of order n for E</u> .

We note that Chebyshev points of order n for E are the zeros of a Chebyshev polynomial of degree *n for* E. *We have claimed that if*

$E$ is of cardinality $k \geq n$, then for $n = 1,2,\ldots,k$ Chebyshev points of order $n$ for $E$ are unique. We shall assume for the rest of the chapter that the cardinality of $E$ is infinite. Thus, for every $n$, it makes sense to refer to the Chebyshev points of order $n$ for $E$. The Chebyshev points for $E$ do not necessarily lie in $E$. However, it is true that:

Theorem 1.2.3 (Convex Hull Theorem)

If $z_1^*, z_2^*, \ldots, z_n^*$ are the Chebyshev points of order $n$ for $E$, then $z_1^*, z_2^*, \ldots, z_n^*$ lie in the convex hull of $E$.

Proof: Consider the following lemma:

Lemma 1.2.1 (Fejér's Principle)

Let $K$ be a compact convex set in $\mathbb{C}$ and $z^* \notin K$. Then there exists a point $\zeta$ such that $|z-\zeta| < |z-z^*|$ for every $z$ in $K$.

Proof: There exists a line $\ell$ strictly separating $z^*$ and $K$. Let $\ell'$ be the line through $z^*$, perpendicular to $\ell$ and let $\zeta$ be the point of intersection of $\ell$ and $\ell'$. Then angle $zz^*\zeta < \frac{1}{2}\pi$ and angle $z\zeta z^* > \frac{1}{2}\pi$ for every $z$ in $K$. Thus $|z-\zeta| < |z-z^*|$.

Proof of Theorem:

Let $z_1^*, z_2^*, \ldots, z_n^*$ be the Chebyshev points of order $n$ for $E$. We shall designate the closed convex hull of $E$ by $K(E)$. Suppose $z_1^* \notin K(E)$. Then by Fejér's Principle (where $K(E)$ is designated by $K$ and $z_1^*$ by $z^*$), there exists a $\zeta$ such that $|z-\zeta| < |z-z_1^*|$.

Consider a polynomial

$$Q_n(Z) = \left[\prod_{i=2}^{n} (z-z_i^*)\right](z-\zeta) . \quad \|Q_n(z)\| = \max_{z\in E} \prod_{i=2}^{n} |z-z_i^*|\ |z-\zeta| <$$

$\max_{z \in E} \prod_{i=1}^{n} |z-z_i^*| = \|T_n(z;E)\|$ . Thus we can find a polynomial whose norm is smaller than the norm of the Chebyshev polynomial for $E$ . Contradiction! Hence, all the Chebyshev points of order $n$ for $E$ lie in $K(E)$ , the closed convex hull of $E$ .

Corollary 1.2.1

If $E$ is convex, then the Chebyshev points of order $n$ for $E$ lie in $E$ .

Like a Fekete polynomial, a Chebyshev polynomial for $E$ also defines a lemniscatic region.

Definition 1.2.4

By $\mathcal{L}_{T_n}$ we shall mean $\{z \,\big|\, |T_n(z;E)| \le m_n(E)\}$ .

One can show:

Theorem 1.2.4

a) $E \subset \mathcal{L}_{T_n}$ and

b) $|\partial\mathcal{L}_{T_n} \cap E| \ge n+1$ .

## 1.3 Fekete's Theorem

We now come to the main theorem of this chapter, Fekete's Theorem. Fekete [6] proved that the transfinite diameter of $E$ and the Chebyshev constant of $E$ are equal.

Theorem 1.3.1

If $E$ is a compact set in the complex plane, then $\delta(E) = \chi(E)$ .

Proof: We note first that if $E$ has only a finite number of points, then both $\delta(E)$ and $\chi(E)$ are zero, so we shall assume that the cardinality of E is infinite.

a) To show $\chi(E) \le \delta(E)$ :

Choose $\xi_1, \xi_2, \ldots, \xi_n$ such that $V_n(E) = V(\xi_1, \xi_2, \ldots, \xi_n)$ . Since

$m_n(E) = \min\limits_{\substack{z_i \in \mathbb{C} \\ i=1,2,\ldots,n}} \max\limits_{z \in E} \prod\limits_{i=1}^{n} |z-z_i|$ we choose z in E such that

$m_n(E) \le \prod\limits_{i=1}^{n} |z-\xi_i|$ . Now

$$V_{n+1}(E) \ge V(z, \xi_1, \xi_2, \ldots, \xi_n) \ge m_n(E) V(\xi_1, \xi_2, \ldots, \xi_n) = m_n(E) V_n(E) \tag{1.3.1}$$

$$d_{n+1}(E) = \Big[V_{n+1}(E)\Big]^{\frac{2}{n(n+1)}} \ge \Big[m_n(E) V_n(E)\Big]^{\frac{2}{n(n+1)}} =$$

$$\{\underbrace{\tau_n \tau_n \cdots \tau_n}_{n \text{ terms}} \; \underbrace{d_n d_n \cdots d_n}_{\binom{n}{2} \text{ terms}}\}^{\frac{2}{n(n+1)}}$$

On the right we have the geometric average of $\binom{n+1}{2}$ terms. Since any average is greater than or equal to its minimal term we have

$$d_{n+1}(E) \ge \min(\tau_n(E), d_n(E)) \; . \tag{1.3.2}$$

We have equality if and only if $\tau_n(E) = d_n(E)$ . But this implies $\chi(E) = \delta(E)$. Thus for every n , either $\tau_n(E) = d_n(E)$ or $d_{n+1}(E) > \min(\tau_n(E), d_n(E))$ . If $d_n(E) < \tau_n(E)$ then $d_{n+1}(E) > d_n(E)$ , which is impossible since Theorem 1.1.1 avers that $\{d_n(E)\}$ is monotone decreasing. Thus $\tau_n(E) < d_n(E)$ . Hence in all cases $\tau_n(E) \le d_n(E)$ and $\chi(E) \le \delta(E)$ .

b) To show $\delta(E) \le \chi(E)$ :

Let $\Delta(z_1,z_2,\ldots,z_{n+1}) = \det|z_i^n, z_i^{n-1},\ldots z_i^2, z_i, 1|$

$$= \begin{vmatrix} z_1^n & z_2^n & \cdots & z_{n+1}^n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ z_1^2 & z_2^2 & \cdots & z_{n+1}^2 \\ z_1 & z_2 & \cdots & z_{n+1} \\ 1 & 1 & \cdots & 1 \end{vmatrix} \qquad (1.3.3)$$

$\Delta(z_1,z_2,\ldots,z_{n+1})$ is called the <u>Vandermonde determinant</u> of the values $z_1,z_2,\ldots,z_{n+1}$ . We note that $|\Delta(z_1,z_2,\ldots,z_{n+1})| = V(z_1,z_2,\ldots,z_{n+1})$. Let $T_n(z;E) = z^n + c_1 z^{n-1} + c_2 z^{n-1} + \ldots + c_n$ be the Chebyshev polynomial of degree $n$ for $E$ . Now, the value of $\Delta(z_1,z_2,\ldots,z_{n+1})$ will not change if we add to the top row multiples of the other rows. We shall choose as our multiples the coefficients of the powers of $z$ in the Chebyshev polynomial for $E$ ; i.e., we multiply the second row by $c_1$ , the third row by $c_2$ , ... up to the $n+1$'st row by $c_n$ and add to the top row. We then have

$$\Delta(z_1,z_2,\ldots,z_{n+1}) = \begin{vmatrix} T_n(z_1;E) & T_n(z_2;E) & \cdots & T_n(z_{n+1};E) \\ z_1^{n-1} & z_2^{n-1} & \cdots & z_{n+1}^{n-1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ z_1^2 & z_2^2 & \cdots & z_{n+1}^2 \\ z_1 & z_2 & \cdots & z_{n+1} \\ 1 & 1 & \cdots & 1 \end{vmatrix} \qquad (1.3.4)$$

Expanding the determinant using elements of the first row gives:

$$\Delta(z_1,z_2,\ldots,z_{n+1}) = \sum_{i=1}^{n+1} (-1)^{i+1}\, T_n(z_i;E)\, \Delta(z_1,z_2,\ldots,\hat{z}_i,\ldots,z_{n+1}) \tag{1.3.5}$$

where $\Delta(z_1,z_2,\ldots,\hat{z}_i,\ldots,z_{n+1})$ indicates the Vandermonde determinant with the column involving $z_i$ and its powers omitted. Thus

$$|\Delta(z_1,z_2,\ldots,z_{n+1})| \le \sum_{i=1}^{n+1} |T_n(z_i;E)|\; |\Delta(z_1,z_2,\ldots,\hat{z}_i,\ldots,z_{n+1})|$$

or $$V(z_1,z_2,\ldots,z_{n+1}) \le \sum_{i=1}^{n+1} |T_n(z_i;E)|\; V(z_1,z_2,\ldots,\hat{z}_i,\ldots,z_{n+1}) .$$

Taking the maximum of both sides we have

$$V_{n+1}(E) \le m_n(E)\; V_n(E)\; (n+1) \tag{1.3.6}$$

Using equations (1.3.1) and (1.3.6), and simplfying the notation by omitting the $E$'s we have

$$m_n V_n \le V_{n+1} \le m_n V_n (n+1) \tag{1.3.7}$$

Now if $\delta(E) = 0$ our conclusion is satisfied automatically, so we assume $\delta(E) > 0$ . Consider the series $\sum_{n=2}^{\infty} V_n^{\frac{1}{n}} z^n$ . Let the radius of convergence be $\rho$ . By the Cauchy-Hademard formula we have

$$\frac{1}{\rho} = \overline{\lim_{n\to\infty}}\; V_n^{\frac{1}{n^2}} \tag{1.3.8}$$

But $$V_n^{\frac{1}{n^2}} = V_n^{\frac{1}{n(n-1)}}\; V_n^{\frac{1}{n^2} - \frac{1}{n(n-1)}} = V_n^{\frac{1}{n(n-1)}}\; V_n^{\frac{-1}{n^2(n-1)}} = \frac{V_n^{\frac{1}{n(n-1)}}}{\left[V_n^{\frac{1}{n(n-1)}}\right]^{\frac{1}{n}}}$$

As $n \to \infty$, the numerator approaches $[\delta(E)]^{\frac{1}{2}}$ and the denominator approaches 1. Thus $V_n^{\frac{1}{n^2}}$ approaches $[\delta(E)]^{\frac{1}{2}}$ and (1.3.8) gives

$$\rho = \frac{1}{[\delta(E)]^{\frac{1}{2}}} \tag{1.3.9}$$

Employing the ratio test for convergence we examine $\lim\limits_{n\to\infty} \dfrac{V_{n+1}^{\frac{1}{n+1}}}{V_n^{\frac{1}{n}}}$.

If this limit exists, it too must be equal to $\frac{1}{\rho}$.

Now $\dfrac{V_{n+1}^{\frac{1}{n+1}}}{V_n^{\frac{1}{n}}} = \left(\dfrac{V_{n+1}}{V_n}\right)^{\frac{1}{n}} V_{n+1}^{\frac{1}{n+1} - \frac{1}{n}} = \left(\dfrac{V_{n+1}}{V_n}\right)^{\frac{1}{n}} \dfrac{1}{V_{n+1}^{\frac{1}{n(n+1)}}}$

By (1.3.7) $\left(\dfrac{V_{n+1}}{V_n}\right)^{\frac{1}{n}}$ approaches $\chi(E)$ as $n \to \infty$, and $\dfrac{1}{V_{n+1}^{\frac{1}{n(n+1)}}}$ approaches $\dfrac{1}{[\delta(E)]^{\frac{1}{2}}}$ as $n \to \infty$. Hence $\dfrac{1}{\rho} = \dfrac{\chi(E)}{[\delta(E)]^{\frac{1}{2}}}$. Using $\rho = \dfrac{1}{[\delta(E)]^{\frac{1}{2}}}$ from (1.3.9) above, we have $\delta(E) = \chi(E)$.

## 1.4 Examples

The following examples will illustrate some of the previous material. We first consider

### Example 1.4.1

Let $E$ be the unit disk in $\mathbb{C}$, which we shall call $D : \{z \,|\, |z| \le 1\}$.

We ask the question, "Where should the $z_i$'s be placed so as to minimize the expression $\max_{z\in D} \prod_{i=1}^{n} |z-z_i|$ ?" Intuitively one thinks placing the $z_i$'s at the center would accomplish this, and indeed it is the case, but let's prove it.

Theorem 1.4.1

$\chi(D) = 1$ .

Proof: Placing $z_i = 0$ for $i=1,2,\dots,n$ we have $m_n(D) \le \max_{z\in D} \prod_{i=1}^{n} |z-0| = \max_{z\in D} |z^n| = 1$ . Suppose $m_n(D) < 1$. Then $\|T_n(z;D)\| < 1$ . Consider $Q_n(z) = z^n T_n\left(\frac{1}{z};D\right)$ . On $\partial D$ we have $|Q_n(z)| = |z^n|\left|T_n\left(\frac{1}{z};D\right)\right| < 1$ and by the maximum modulus principle $|Q_n(z)| < 1$ throughout $D$ . But at $z = 0$ , $Q_n(0) = 1$ . Contradiction! Hence for every $n$ , $m_n(D) = 1$ , $\tau_n(D) = 1$ , and $\chi(D) = 1$ .

Corollary 1.4.1

$\chi(D_R) = R$ , where $D_R$ is a disk of radius $R$ .

Thus for the unit disk, $T_n(z) = z^n$ and the Chebyshev points of order $n$ for $D$ for all $n$ , are located at $z = 0$ . From Theorem 1.3.1 we know that $\delta(D) = 1$ . It can be shown that a set of Fekete points of order $n$ for $D$ are located at the $n$'th roots of unity. Unlike the set of Chebyshev points for the disk, this set is not unique, for any set of points which are the $n$'th roots of $\omega$ where $|\omega| = 1$ has the property characterizing Fekete points.

Example 1.4.2

Now let us consider $E$ to be the interval of the real line $[-1,1]$

which we shall denote by $I$ . We again ask the same question. How should the points $x_1, x_2, \ldots, x_n$ be placed so as to minimize the expression $\max\limits_{x \in I} \prod\limits_{i=1}^{n} |x-x_i|$ ? The answer here is not at all obvious. For reasons which will become apparent later we consider the polynomial

$$C_n(x) = \cos n\theta \quad \text{where} \quad \cos\theta = x \tag{1.4.1}$$

We note that $C_0(x) = 1$ and $C_1(x) = x$ . Using trigonometric identities it is easy to show that the polynomials $C_n(x)$ satisfy the three term recurrence relation:

$$C_{n+1}(x) = 2xC_n(x) - C_{n-1}(x) \tag{1.4.2}$$

Thus

$$C_{n+1}(x) = 2^{n-1}x^n + \text{terms of lower degree} \tag{1.4.3}$$

Although $C_n(x)$ is referred to in the literature as the Chebyshev polynomial of degree $n$ for $I$ , and usually is denoted by $T_n(x)$ , we have reserved that term and that notation for a monic polynomial. Hence we shall look at $\hat{C}_n(x)$ where

$$\hat{C}_n(x) = \frac{1}{2^{n-1}} C_n(x) \tag{1.4.4}$$

and we shall show that $\hat{C}_n(x) = \frac{1}{2^{n-1}} \cos(\, n \arccos x \,)$ is indeed the unique Chebyshev polynomial $T_n(x;I)$ of degree $n$ for $I$ . But first let us look at the zeros of $T_n(x;I)$ which are the Chebyshev points of order $n$ for $I$ .

**Theorem 1.4.2**

$T_n(x;I)$ has simple zeros at the $n$ points $x_k{}^* = \cos\left(\frac{2k-1}{2n}\right)\pi$ $k = 1,2,\ldots,n$ . On $[-1,1]$ $T_n(x)$ has extreme values at the $n+1$ points $\bar{x}_k = \cos \frac{k}{n}\pi$ $k = 0,1,2,\ldots,n$ where it assumes the value $(-1)^k \frac{1}{2^{n-1}}$ .

**Proof:** See Davis [ 4 ].

**Theorem 1.4.3**

$T_n(x;I) = \frac{1}{2^{n-1}} \cos ( n \operatorname{arc} \cos x )$ is the Chebyshev polynomial for $I$ and $\chi(I) = \frac{1}{2}$ .

**Proof:**

We have $m_n(I) \leq \|T_n(x;I)\| = \max\limits_{x\in I} |T_n(x;I)| = \frac{1}{2^{n-1}}$ . Suppose $m_n(I) < \frac{1}{2^{n-1}}$ . Let $Q_n(x)$ be a monic polynomial such that $\|Q_n(x)\| < \frac{1}{2^{n-1}}$ . Consider $P(x) = T_n(x;I) - Q_n(x)$ . Now at $\bar{x}_k$ $P(\bar{x}_k) = \frac{1}{2^{n-1}} (-1)^k - Q_n(\bar{x}_k)$ and since $|Q_n(\bar{x}_k)| < \frac{1}{2^{n-1}}$ , $P(\bar{x}_k)$ takes on alternately $+$ and $-$ values for the $n+1$ values $k = 0,1,2,\ldots,n$ . Thus $P(x)$ has $n$ zeros. But since both $T_n(x;I)$ and $Q_n(x)$ are monic, $P(x)$ is of degree $n-1$ . Hence $P(x) \equiv 0$ and $Q_n(x) \equiv T_n(x;I)$ . Now $\frac{1}{2^{n-1}} = \max\limits_{x\in I} |T_n(x;I)| = \max\limits_{x\in I} |Q_n(x)| < \frac{1}{2^{n-1}}$ . Contradiction! Thus $m_n(I) = \frac{1}{2^{n-1}}$ and $\chi(I) = \lim\limits_{n\to\infty} \tau_n(I) = \lim\limits_{n\to\infty}\left(\frac{1}{2^{n-1}}\right)^{\frac{1}{n}} = \frac{1}{2}$ .

From (1.1.7) and Theorem 1.3.1, for the interval $\mathcal{J} = [a,b]$ $a \leq b$ we have

$$\chi(\mathcal{J}) = \tfrac{1}{4}(b - a) \qquad (1.4.5)$$

Again we know that $\delta(I) = \frac{1}{2}$, but what can one say about the Fekete points for $I$? Without getting involved in a discussion of orthogonal polynomials, we shall state that the Fekete points of order $n$ for $I$ are the points $\pm 1$ and the zeros of the derivative of the Legendre polynomial of degree $n-1$. (Refer to Szego [31] p. 379, prob. 37 for clarification.) Unlike the case for the disk, the Fekete points for the interval are unique.

## 1.5 Additional Observations

The above examples were chosen since direct calculation is not too involved. In most cases we rely on the following mapping theorem to calculate the transfinite diameter for a larger class of compact sets $E$.

### Theorem 1.5.1

Let $E$ be a bounded continuum with connected complement. If $E^c_\infty$ (see Def. 1.1.5) is mapped conformally on $|w| > R$ by $w = w(z) = z + a_0 + \frac{a_1}{z} + \frac{a_2}{z^2} + \cdots$ then the radius $R$, called the outer mapping radius of $E$, is equal to $\delta(E)$.

Proof: See Hille [14].

Although we shall not deal with it again, the classical theory of this subject would not be complete if one did not include the following definition:

Let $E$ be a compact set in $\mathbb{C}$ and $\mu \geq 0$ be a positive mass distribution on E of total mass 1 . We consider the energy integral

$$I(\mu) = \iint_E \log \frac{1}{|a-b|} \, d\mu(a) \, d\mu(b) \qquad (1.5.1)$$

Let

$$V = \inf_{\mu} I(\mu) \qquad (1.5.2)$$

It can be shown $-\infty < V \le \infty$ .

<u>Definition 1.5.1</u>

Let $\gamma(E) = e^{-V}$ . We call $\gamma(E)$ <u>the capacity of E</u> .

The notion of capacity is part of our story because of a result by Szegö [30]. He proved that

$$\gamma(E) = \delta(E) \qquad (1.5.3)$$

Another definition, one that we shall refer to occasionally is:

<u>Definition 1.5.2</u>

A monic polynomial of degree n of smallest norm defined on E , <u>whose zeros lie in E</u> , shall be called a <u>restricted Chebyshev polynomial of degree n for E</u> and denoted by $\tilde{T}_n(z;E)$ .

We let

$$\tilde{m}_n(E) = \|\tilde{T}_n(z;E)\|_E \qquad (1.5.4)$$

and

$$\tilde{\tau}_n(E) = \left[\tilde{m}_n(E)\right]^{\frac{1}{n}} \qquad (1.5.5)$$

One can show $\lim_{n\to\infty} \tilde{\tau}_n(E)$ exists (see Goluzin [10]) and

Definition 1.5.3

$\tilde{\chi}(E) = \lim_{n\to\infty} \tilde{\tau}_n(E)$ is called the restricted Chebyshev constant for E .

Theorem 1.5.3

$\tilde{\chi}(E) = \chi(E)$

Proof: Since the minimum is taken over a larger set, $m_n(E) \le \tilde{m}_n(E)$ , thus $\chi(E) \le \tilde{\chi}(E)$ . Now, a closer examination of Theorem 1.3.1 reveals that in part a) we proved $\tilde{\tau}_n(E) \le d_n(E)$ . Thus $\tilde{\chi}(E) \le \delta(E)$ . But in part b) we proved $\delta(E) = \chi(E)$ . Hence, $\tilde{\chi}(E) = \chi(E)$ .

$\tilde{T}_n(z;E)$ is called a restricted Chebyshev polynomial of degree n for E , since we do not necessarily have uniqueness in this case, although the cardinality of E may be infinite.

Definition 1.5.4

The zeros of $\tilde{T}_n(z;E)$ will be called restricted Chebyshev points of order n for E .

If E consists of more than one point and is convex, it can be shown that for every n , the restricted Chebyshev polynomial of degree n for E is unique and is the same as the Chebyshev polynomial of degree n for E . However, let us consider:

Example 1.5.1

Let E be $\partial D$ , i.e., the unit circle. The Chebyshev points of order n for $\partial D$ are $z_i = 0$ $i = 1,2,\ldots,n$ , as in the case for D . A set of Fekete points of order n for $\partial D$ are the n'th roots of $\omega$ where $|\omega| = 1$ , as in the case for D . Let us look at this example in the restricted case.

Lemma 1.5.1

If $q_n(z)$ is a polynomial of degree $n$ and has its zeros on $|z| \geq 1$, then

$$\max_{|z|=R>1} |q_n(z)| \leq \frac{1+R^n}{2} \max_{|z|=1} |q_n(z)|$$

Proof: See Ankeny and Rivlin [1].

Theorem 1.5.4

$\tilde{T}_n(z;\partial D) = z^n - \omega$ where $|\omega| = 1$.

Proof: $\|\tilde{T}_n(z;\partial D)\| = \max_{z \in \partial D} |z^n - \omega| = 2$. Thus $\tilde{m}_n(\partial D) \leq 2$. Suppose there exists a monic polynomial $q_n(z)$ such that $\|q_n(z)\| = \rho < 2$, and $q_n(z)$ has its zeros on $\partial D$. Then by the preceding lemma we have $\max_{|z|=R>1} |q_n(z)| \leq \frac{1+R^n}{2} \rho$. Let $\theta = \frac{\rho}{2}$. Then $\theta < 1$. For $R$ large enough $(1+R^n)\theta < R^n$. Thus $\max_{|z|=R>1} |q_n(z)| < R^n = \max_{|z|=R} |z^n|$. We apply Rouche's Theorem. Hence $z^n - q_n(z)$ has the same number of zeros as $z^n$ on $|z| \leq R$. But $z^n - q_n(z)$ is a polynomial of degree at most $n-1$, whereas $z^n$ is a polynomial of degree $n$. Thus $q_n(z) - z^n \equiv 0$ and $q_n(z) \equiv z^n$. But $z^n$ has no zeros on $\partial D$. Contradiction! Thus there does not exist such a polynomial $q_n(z)$ and $\tilde{T}_n(z;\partial D) = z^n - \omega$ where $|\omega| = 1$ is a restricted Chebyshev polynomial of degree $n$ for $\partial D$.

In this case we see that the set of restricted Chebyshev polynomials for $\partial D$ is the same as the set of Fekete polynomials for $\partial D$, and the sets of restricted Chebyshev points and Fekete points coincide.

Our summary of the classical theory is complete. Other investigations in the complex plane were carried out by Frostman [7] and Tsuji [32] using spherical, elliptical, and hyperbolic metrics. Their definitions of transfinite diameter, Chebyshev constant, and capacity using these metrics led to the equality of all three in the complex plane for any compact set $E$.

# CHAPTER 2

## Extensions and Generalizations

Fekete's definitions of the transfinite diameter and the Chebyshev constant involve the choice of a metric space (complex plane with usual metric), a subset (compact set), and an averaging process (geometric mean). In Chapter 1 we mentioned the work of those who dealt with other metrics in the complex plane. In this chapter we will generalize by considering other averaging processes, and extend the definitions to any metric space.

### 2.1 Averaging Processes

The geometric mean belongs to a class of averages which we shall now discuss. Let $x_1, x_2, \ldots, x_n$ be a set of non-negative numbers.

Definition 2.1.1

The function $A_{(\lambda)}(x_1, x_2, \ldots, x_n) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^{\lambda}\right)^{\frac{1}{\lambda}}$ , $\lambda \neq 0$ will be called the λ-th power average of $x_1, x_2, \ldots, x_n$ . If for some i , $i = 1, 2, \ldots, n$ $x_i = 0$ and $\lambda < 0$ then we define $A_{(\lambda)}(x_1, x_2, \ldots, x_n) = 0$ .

Definition 2.1.2

For $\lambda = 1$ , $A_{(1)}(x_1, x_2, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i$ is called the arithmetic average of $x_1, x_2, \ldots, x_n$ .

Definition 2.1.3

For $\lambda = -1$ and $x_i \neq 0$ , $i = 1,2,\ldots,n$ , $A_{(-1)}(x_1,x_2,\ldots,x_n) =$

$$\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$$ is called the harmonic average of $x_1,x_2,\ldots,x_n$ .

Theorem 2.1.1

$$\lim_{\lambda\to 0} A_{(\lambda)}(x_1,x_2,\ldots,x_n) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$$

Proof: See Hardy, Littlewood, and Pólya [11].

Definition 2.1.4

We define $A_{(0)}(x_1,x_2,\ldots,x_n) = \lim_{\lambda\to 0} A_{(\lambda)}(x_1,x_2,\ldots,x_n)$ . Thus,

$$A_{(0)}(x_1,x_2,\ldots,x_n) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$$ which we call the geometric average[1] of $x_1,x_2,\ldots,x_n$ .

The $\lambda$-th power averages have the following properties, proofs of which may be found in the previous reference cited.

Theorem 2.1.2

a) If $\lambda < \mu$ , then $A_{(\lambda)}(x_1,x_2,\ldots,x_n) < A_{(\mu)}(x_1,x_2,\ldots,x_n)$ unless $x_1=x_2=\ldots=x_n$, or $\mu \leq 0$ and $x_i = 0$ for some $i$ , $i = 1,2,\ldots,n$ .

b) If $\frac{1}{\lambda} + \frac{1}{\mu} = 1$ and $\lambda > 1$ , then

---

[1] We have referred to this previously as the geometric mean. We shall use the words "mean" and "average" interchangeably.

$\frac{1}{n}\sum_{i=1}^{n} x_i y_i \leq A_{(\lambda)}(x_1,x_2,\ldots,x_n)A_{(\mu)}(y_1,y_2,\ldots,y_n)$ with equality if and only if the set $\left\{x_1^{\lambda},x_2^{\lambda},\ldots,x_n^{\lambda}\right\}$ is proportional to the set $\left\{y_1^{\mu},y_2^{\mu},\ldots,y_n^{\mu}\right\}$ or either $x_1=x_2=\ldots x_n=0$ or $y_1=y_2=\ldots=y_n=0$ . (Holder's Inequality.)

c) For $\lambda \geq 1$ , $A_{(\lambda)}(x_1+y_1,x_2+y_2,\ldots,x_n+y_n)$

$\leq A_{(\lambda)}(x_1,x_2,\ldots,x_n) + A_{(\lambda)}(y_1,y_2,\ldots,y_n)$

with equality if and only if $\lambda = 1$ or the set $\{x_1,x_2,\ldots,x_n\}$ is proportional to the set $\{y_1,y_2,\ldots,y_n\}$ . (Minkowski's Inequality.)

Since it will be used in the sequel, we define the concept of weighted average.

Definition 2.1.5

Let $w_i > 0$ $i = 1,2,\ldots,n$ . The function

$$A_{(\lambda)}(x_1,x_2,\ldots,x_n;w_1,w_2,\ldots,w_n) = \left[\frac{\sum_{i=1}^{n} w_i x_i^{\lambda}}{\sum_{i=1}^{n} w_i}\right]^{\frac{1}{\lambda}}$$

$\lambda \neq 0$ will be called the weighted $\lambda$-th power average of $x_1,x_2,\ldots,x_n$ , with respect to the weights $w_1,w_2,\ldots,w_n$ . If for some $i$ , $x_i = 0$ and $\lambda < 0$ then we set $A_{(\lambda)}(x_1,x_2,\ldots,x_n;w_1,w_2,\ldots,w_n) = 0$ .

Definition 2.1.6

$$A_{(0)}(x_1,x_2,\ldots,x_n;w_1,w_2,\ldots,w_n) = \left[\prod_{i=1}^{n} x_i^{w_i}\right]^{\frac{1}{\sum_{i=1}^{n} w_i}}$$

The weighted $\lambda$-th power average reduces to the ordinary $\lambda$-th power average when $w_i = 1$ for all $i = 1,2,\ldots,n$ .

Let $E$ be a compact set in a metric space $(X,\rho)$ . Using the $\lambda$-th power means one can consider:

$$d_n^{\lambda}(E) = \max_{\substack{x_i \in E \\ i=1,2,\ldots,n}} \left[\frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} \rho(x_i,x_j)^{\lambda}\right]^{\frac{1}{\lambda}} \tag{2.1.1}$$

and

$$\tau_n^{(\lambda)}(E) = \min_{\substack{x_i \in X \\ i=1,2,\ldots,n}} \max_{x \in E} \left[\frac{1}{n} \sum_{i=1}^{n} \rho(x,x_i)^{\lambda}\right]^{\frac{1}{\lambda}} \tag{2.1.2}$$

One can show that both $\lim\limits_{n\to\infty} d_n^{(\lambda)}(E)$ and $\lim\limits_{n\to\infty} \tau_n^{(\lambda)}(E)$ exist.

Definition 2.1.7

$\delta^{(\lambda)}(E) = \lim\limits_{n\to\infty} d_n^{(\lambda)}(E)$ is called the <u>$\lambda$-th power average transfinite diameter for E</u> .

Definition 2.1.8

$\chi^{(\lambda)}(E) = \lim\limits_{n\to\infty} \tau^{(\lambda)}(E)$ is called the <u>$\lambda$-th power average Chebyshev constant for E</u> .

Pólya and Szegö [25] considered the sets $I$, $\partial I$, $D$, and $\partial D$ in the complex plane using $\lambda$-th power averages. In addition they found values for the transfinite diameter and the Chebyshev constant for the unit ball and the unit sphere in three dimensional Euclidean space. A table of their results may be found in the appendix.

The arithmetic, geometric, and harmonic means were also used by Leja in a series of papers on this topic dating back to 1933. Leja "weakened" the metric in the plane by his use of a function $\omega(z_i, z_j)$ that obeyed only the laws of positivity and symmetry for a metric function, without the triangle law. Leja also examined various associated sequences of polynomials, extension of the problem to two dimensional complex space, and some conformal mapping aspects. See Leja [20],[21],[22].

Now let us consider a class of averages which are based on the following postulates imposed by Kolmogorov [18] and also by Nagumo [23].

## THE POSTULATES

a) For each natural number $n$ and for every set of $n$ positive values $x_1, x_2, \ldots, x_n$, there exists a positive average $A(x_1, x_2, \ldots, x_n)$.

b) $A(x_1, x_2, \ldots, x_n)$ is a continuous symmetric function of its arguments and $A(x_1, x_2, \ldots, x_n)$ is strictly increasing as a function in each of them.

c) $A(x, x, \ldots, x) = x$.

d) $A(x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_n) = A(y, y, \ldots, y, x_{k+1}, \ldots, x_n)$ if $y = A(x_1, x_2, \ldots, x_k)$.

As a consequence of POSTULATE b) and POSTULATE c) we have the basic inequality:

$$\min_i x_i \le A(x_1,x_2,\ldots,x_n) \le \max_i x_i \qquad (2.1.3)$$

with equality holding if and only if $x_1 = x_2 = \ldots = x_n$ .

If the definition of $A_{(\lambda)}(x_1,x_2,\ldots,x_n)$ is restricted to only positive values of its arguments, then the $\lambda$-th power average fulfills the POSTULATES. We have, however, defined $A_{(\lambda)}(x_1,x_2,\ldots,x_n)$ for non-negative values of its arguments, and we can extend the definition of an average $A(x_1,x_2,\ldots,x_n)$ by continuity when one or more of its arguments is zero.

We would expect to have $A(0,0,\ldots,0) = 0$ . However, $A(x_1,x_2,\ldots,x_n)$ can be equal to 0 without having all of its arguments equal to 0 . For example, consider $A_{(\lambda)}(x_1,x_2,\ldots,x_n)$ for $\lambda \le 0$ and $x_i = 0$ for some $i$ , $i = 1,2,\ldots,n$ .

The statement of POSTULATE b) - that $A(x_1,x_2,\ldots,x_n)$ is strictly increasing as a function in each of its arguments - holds only when one considers strictly positive values of the arguments. In view of this, the basic inequality (2.1.3) must be modified if we consider $A(x_1,x_2,\ldots,x_n)$ defined for non-negative arguments. In this case we have

$$\min_i x_i \le A(x_1,x_2,\ldots,x_n) \le \max_i x_i \qquad (2.1.4)$$

with equality on the right if and only if all the $x_i$'s, $i = 1,2,\ldots,n$ are equal, and equality on the left if and only if $A(x_1,x_2,\ldots,x_n) = 0$ when some $x_i = 0$ , $i = 1,2,\ldots,n$ or all the $x_i$'s, $i = 1,2,\ldots,n$ are equal.

Another consequence of the POSTULATES, one that will be used in several proofs, is the following:

Theorem 2.1.3

If $A(y_{k+1},y_{k+2},\ldots,y_n) \geq A(x_{k+1},x_{k+2},\ldots,x_n)$ then $A(x_1,x_2,\ldots,x_k,y_{k+1},y_{k+2},\ldots,y_n) \geq A(x_1,x_2,\ldots,x_k, x_{k+1},x_{k+2},\ldots,x_n)$ .

Proof: Let $A(y_{k+1},y_{k+2},\ldots,y_n) = \gamma$ and $A(x_{k+1},x_{k+2},\ldots,x_n) = \delta$ . Then by POSTULATE d), $A(x_1,x_2,\ldots,x_k,y_{k+1},y_{k+2},\ldots,y_n) = A(x_1,x_2,\ldots,x_k,\gamma,\gamma,\ldots,\gamma)$ and $A(x_1,x_2,\ldots,x_k,x_{k+1},x_{k+2},\ldots,x_n) = A(x_1,x_2,\ldots,x_k,\delta,\delta,\ldots,\delta)$ . By POSTULATE b), since $\gamma \geq \delta$ , the conclusion follows.

Let $F(t)$ be a real-valued function, which is continuous and strictly monotone for $0 < t < \infty$ . The function $F(t)$ will generate a function $A_F(x_1,x_2,\ldots,x_n)$ by means of the correspondence:

$$A_F(x_1,x_2,\ldots,x_n) = F^{-1}\left( \frac{1}{n} \sum_{i=1}^{n} F(x_i)\right) \qquad (2.1.5)$$

It can be shown that the function $A_F(x_1,x_2,\ldots,x_n)$ so generated satisfies the POSTULATES for an average function. Moreover, to every average satisfying the POSTULATES, there corresponds a generating function which is related to the average by means of (2.1.5). See Hille [16].

We note that the function $F(t) = t^\lambda$ generates the $\lambda$-th power average for $\lambda \neq 0$ , and the function $F(t) = \log t$ generates the 0-th power, or geometric average. Indeed the $\lambda$-th power averages are characterized by:

Theorem 2.1.4

If for every $k$ , $A(kx_1,kx_2,\ldots,kx_n) = kA(x_1,x_2,\ldots,x_n)$ , (i.e., $A(x_1,x_2,\ldots,x_n)$ is homogeneous of degree one), then $A(x_1,x_2,\ldots,x_n)$

is a $\lambda$-th power average, and conversely.

The correspondence between $A_F(x_1,x_2,\ldots,x_n)$ and its generating function $F(t)$ is not unique. We have the following theorem:

Theorem 2.1.5

Let $F(t)$ generate $A_F(x_1,x_2,\ldots,x_n)$ and $G(t)$ generate $A_G(x_1,x_2,\ldots,x_n)$ . Then $A_F(x_1,x_2,\ldots,x_n) = A_G(x_1,x_2,\ldots,x_n)$ for every set $\{x_1,x_2,\ldots,x_n\}$ if and only if $F(t) = \alpha G(t) + \beta$ where $\alpha \neq 0$ .

Theorem 2.1.2, part a) gives an order relationship for $\lambda$-th power averages. The question arises whether one has a like theorem in the case of averages satisfying the POSTULATES. The answer is affirmative.

Theorem 2.1.6

Let $F(t)$ generate $A_F(x_1,x_2,\ldots,x_n)$ and $G(t)$ generate $A_G(x_1,x_2,\ldots,x_n)$ . Suppose $F(t)$ is increasing. Then $A_G(x_1,x_2,\ldots,x_n) \leq A_F(x_1,x_2,\ldots,x_n)$ for every set $\{x_1,x_2,\ldots,x_n\}$ if and only if $H(y) = F(G^{-1}(y))$ is convex for all $y$ in the range of $G$ .

The proofs of Theorems 2.1.4, 2.1.5, and 2.1.6 may be found in Hardy, Littlewood, and Pólya [11].

Hille [13],[14],[16] approached the task of generalizing Fekete's work by considering averages satisfying the POSTULATES. We too will consider those averages. Each time we write $A(x_1,x_2,\ldots,x_n)$ we mean an average function satisfying the POSTULATES. All theorems referring to "for every average" mean "for every average satisfying the POSTULATES." The material in Sections 2.2 and 2.3 parallels some of Hille's work, but the approach of working with the generating function for the averages is different.

## 2.2 A Generalized Transfinite Diameter

From Theorem 2.1.5 we see that any average $A_F(x_1,x_2,\ldots,x_n)$ is generated not only by $F(t)$ but by any function belonging to the set

$S = \{\alpha F(t) + \beta | \alpha \neq 0\}$ . Consider the subset $T \subset S$ where $T = \{\alpha F(t) | \alpha \neq 0\}$ . Let $G(t) \in T$ and choose $G(t)$ so that it is strictly increasing. This is, of course, possible since $F(t)$ is monotone. Define a function

$$R(u_1,u_2,\ldots,u_n) = G^{-1}\left[nG(A_G(u_1,u_2,\ldots,u_n))\right] \tag{2.2.1}$$

where $G \in T$ .

(Note that we consider $A(u_1,u_2,\ldots,u_n)$ here only in its non-extended sense, that is for only <u>positive</u> values of its arguments. Hence $A_G(u_1,u_2,\ldots,u_n) > 0$ and thus $G(A_G(u_1,u_2,\ldots,u_n))$ is defined.)

<u>Theorem 2.2.1</u>

The function $R(u_1,u_2,\ldots,u_n)$ depends only on the subset $T$ and is independent of the choice of generating function from $T$ .

<u>Proof</u>: Let $R(u_1,u_2,\ldots,u_n)$ be defined as in (2.2.1) and let $Q(u_1,u_2,\ldots,u_n)$ be defined by $Q(u_1,u_2,\ldots,u_n) = H^{-1}\left[nH(A_H(u_1,u_2,\ldots,u_n))\right]$ where $H \in T$ . Since both $G$ and $H$ are in $T$ , we know $A_G(u_1,u_2,\ldots,u_n)$ is equal to $A_H(u_1,u_2,\ldots,u_n)$ . Moreover, for some $\gamma \neq 0$ , $G(t) = \gamma H(t)$ . Thus $R(u_1,u_2,\ldots,u_n) = G^{-1}\left[nG(A_G(u_1,u_2,\ldots,u_n)\right]$

$$= G^{-1}\left[n\gamma H(A_H(u_1,u_2,\ldots,u_n))\right] = H^{-1}\left[\frac{n\gamma H(A_H(u_1,u_2,\ldots,u_n))}{\gamma}\right]$$

$$= H^{-1}\left[nH(A_H(u_1,u_2,\ldots,u_n))\right] = Q(u_1,u_2,\ldots,u_n) .$$

Let $(X,\rho)$ be a metric space and $E$ a compact subset of $X$ of infinite cardinality. For $n \geq 2$ let $x_1,x_2,\ldots,x_n$ be $n$ distinct points of $E$ . Let $A_G(u_1,u_2,\ldots,u_n)$ be generated by $G \in T$ with $G(t)$ increasing. Define $V^{A(T)}(x_1,x_2,\ldots,x_n)$ by:

$$V^{A(T)}(x_1,x_2,\ldots,x_n) = G^{-1}\left[\binom{n}{2} G(A_G(\rho(x_i,x_j)))\right] \tag{2.2.2}$$

where $\rho(x_i,x_j)$ denotes the $\binom{n}{2}$ distances for $1 \le i < j \le n$ .

To simplify the notation we shall write $V^A(x_1,x_2,\ldots,x_n)$ rather than $V^{A(T)}(x_1,x_2,\ldots,x_n)$ with the understanding that the average A is generated from a function from the subset T .

Since both $V^A(x_1,x_2,\ldots,x_n)$ and $A_G(\rho(x_i,x_j))$ are continuous functions defined on the compact set E , they attain their maximum on E . Set

$$V_n^A(E) = \max_{\substack{x_i \in E \\ i=1,2,\ldots,n}} V^A(x_1,x_2,\ldots,x_n) \tag{2.2.3}$$

and

$$d_n^A(E) = \max A_G(\rho(x_i,x_j)) \tag{2.2.4}$$

where the maximum of $A_G(\rho(x_i,x_j))$ is taken as the n points range over E .[1]

Theorem 2.2.2

$$V_n^A(E) = G^{-1}\left[\binom{n}{2} G(d_n^A(E))\right]$$

Proof: Omitting the arguments for brevity, one has from equation (2.2.2)

---

[1] The set E may be taken to be merely bounded. In this case we can define $d_n^A(E) = \sup A_G(\rho(x_i,x_j))$ , where the supremum is taken as the n points range over E .

$G(V^A) = \binom{n}{2}G(A_G)$ . Hence $\max G(V^A) = \max \binom{n}{2}G(A_G)$ where the maximum is taken over the points $x_1, x_2, \ldots, x_n \in E$ . Since $G$ was chosen increasing we have $G(\max V^A) = \binom{n}{2}G(\max A_G)$ or $G(V_n^A(E)) = \binom{n}{2}G(d_n^A(E))$ . Taking the inverse of both sides gives the desired result.

We note that when $G(t) = \log t$ , then $[V_n(E)]^{\frac{1}{\binom{n}{2}}} = d_n(E)$ and this corresponds with the classical case.

Definition 2.2.1

A set of points $\xi_1, \xi_2, \ldots, \xi_n$ such that $V_n^A(E) = V^A(\xi_1, \xi_2, \ldots, \xi_n)$ shall be called generalized Fekete points of order $n$ for $E$ .

Theorem 2.2.3

The sequence $\{d_n^A(E)\}$ is monotone decreasing.

Proof: Let $\xi_1, \xi_2, \ldots, \xi_{n+1}$ be a set of generalized Fekete points of order $n+1$ for $E$ . Then

$$V_{n+1}^A(E) = V^A(\xi_1, \xi_2, \ldots, \xi_{n+1}) = G^{-1}\left[\binom{n+1}{2}G(A_G(\rho(\xi_i, \xi_j)))\right]$$

for $1 \leq i < j \leq n+1$ . We can write out the arguments for the average by first exhibiting only those that contain $\xi_1$ , thusly:

$$V_{n+1}^A(E) = G^{-1}\left[\binom{n+1}{2}G(A_G(\rho(\xi_1, \xi_2), \rho(\xi_1, \xi_3), \ldots, \rho(\xi_1, \xi_{n+1}), \rho(\xi_i, \xi_j)))\right]$$

where $\rho(\xi_i, \xi_j)$ stands for all pairs $2 \leq i < j \leq n+1$ . If $\zeta_1, \zeta_2, \ldots, \zeta_n$ are generalized Fekete points of order $n$ for $E$ then from Theorem 2.1.2 we have

$$V_{n+1}^A(E) \leq G^{-1}\left[\binom{n+1}{2}G(A_G(\rho(\xi_1, \xi_2), \rho(\xi_1, \xi_3), \ldots, \rho(\xi_1, \xi_{n+1}), \rho(\zeta_i, \zeta_j)))\right]$$

for $1 \le i < j \le n$ . Let $A_1 = A_G(\rho(\xi_1,\xi_2),\rho(\xi_1,\xi_3),\ldots,\rho(\xi_1,\xi_{n+1}),\rho(\zeta_i,\zeta_j))$ .

We write $G^{-1}\left[\frac{1}{\binom{n+1}{2}} G(V^A_{n+1}(E))\right] \le A_1$ .

Similarly $G^{-1}\left[\frac{1}{\binom{n+1}{2}} G(V^A_{n+1}(E))\right] \le A_2$ , where $A_2$ is characterized by having $\xi_2$ as the distinguished point. Likewise we have

$$G^{-1}\left[\frac{1}{\binom{n+1}{2}} G(V^A_{n+1}(E))\right] \le A_i \quad \text{for} \quad i = 3,4,\ldots,n+1 .$$

Taking any average $A$ of the $n+1$ terms on both sides we have on the left by POSTULATE c), $G^{-1}\left[\frac{1}{\binom{n+1}{2}} G(V^A_{n+1}(E))\right]$ and on the right we have $A(A_1,A_2,\ldots,A_{n+1})$ . Examining the arguments in the averages $A_i$ for $i = 1,2,\ldots,n+1$ , one can write

$$A(A_1,A_2,\ldots,A_{n+1}) = A(\underbrace{\rho(\xi_k,\xi_\ell)^2}_{n(n+1)\text{ terms}};\ \underbrace{\rho(\zeta_i,\zeta_j)}_{\binom{n}{2}(n+1)\text{ terms}}) \quad \text{for} \quad 1 \le i < j \le n \quad \text{and}$$

$1 \le k < \ell \le n+1$ . By POSTULATE d), we can replace the arguments on the right by their averages. Thus we have

$$G^{-1}\left[\frac{1}{\binom{n+1}{2}} G(V^A_{n+1}(E))\right] \le A(A_G(\rho(\xi_k,\xi_\ell);A_G(\rho(\zeta_i,\zeta_j)))$$

$$\le A\left(G^{-1}\left[\frac{1}{\binom{n+1}{2}} G(V^A_{n+1}(E))\right];\ G^{-1}\left[\frac{1}{\binom{n}{2}} G(V^A_n(E))\right]\right) .$$

From Theorem 2.2.2 this gives

$$d^A_{n+1}(E) \le A(\underbrace{d^A_{n+1}(E)}_{n(n+1)\text{ terms}} , \underbrace{d^A_n(E)}_{\binom{n}{2}(n+1)\text{ terms}}) .$$

Either $d^A_{n+1}(E) = d^A_n(E)$ or $d^A_{n+1}(E) < d^A_n(E)$ . For if $d^A_{n+1}(E) > d^A_n(E)$ , then the basic inequality for averages (2.1.3) gives $d^A_{n+1}(E) \le A(d^A_{n+1}(E) , d^A_n(E)) < d^A_{n+1}(E)$ . Contradiction! Thus $\{d^A_n(E)\}$ is monotone decreasing.

Definition 2.2.2

$\delta^A(E) = \lim_{n\to\infty} d^A_n(E)$ is called a generalized transfinite diameter of E .

Since $\rho(x_i,x_j) \le d(E)$ for all pairs $(x_i,x_j)$ we have $A(\rho(x_i,x_j)) \le d(E)$ from POSTULATES b) and c). Hence

$$\delta^A(E) \le d(E) . \qquad (2.2.5)$$

Thus, even in the general case, the topological diameter of a set provides an upper bound for the transfinite diameter of that set.

The properties of monotony and continuity of the transfinite diameter hold in the general case. For a proof of a theorem analogous to Theorem 1.1.2, see Hille [16]. From Theorem 2.1.2 part a) for $\lambda$-th power averages we have the result that a generalized transfinite diameter is monotone with respect to $\lambda$ :

$$\text{If } \lambda < \mu \text{ then } \delta^{(\lambda)}(E) \le \delta^{(\mu)}(E) . \qquad (2.2.6)$$

In the general case one does not have statements analogous to

(1.1.7) and (1.1.8). We can say however:

Theorem 2.2.4

Let $(X,\rho)$ be a metric space and let $T$ be a contraction mapping on $(X,\rho)$. If $E$ is a compact subset of $X$, then for any average $A$, $\delta^A(TE) \le \delta^A(E)$, where $TE$ is the image of $E$ under $T$. We have equality if $T$ is an isometry.

Proof: Since $T$ is a contraction mapping, $\rho(Tx_i,Tx_j) < \rho(x_i,x_j)$ for all $x_i,x_j \in E$. Hence $A(\rho(Tx_i,Tx_j)) < A(\rho(x_i,x_j))$ for $1 \le i < j \le n$ by POSTULATE b). Thus $d_n^A(TE) < d_n^A(E)$ which implies $\delta^A(TE) \le \delta^A(E)$. The sufficiency condition for equality is obvious.

Definition 2.2.3

Let $\xi_1,\xi_2,\dots,\xi_n$ be a set of generalized Fekete points of order $n$ for $E$. We consider $F_n^A(x;E) = G^{-1}\left[nG(A_G(\rho(x,\xi_1),\rho(x,\xi_2),\dots,\rho(x,\xi_n)))\right]$ where $G \in T$. The function $F_n^A(x;E)$ will be called a generalized Fekete function of order $n$ for $E$.

We note that $F_n^A(x;E) > 0$. Let

$$\|F_n^A(x;E)\|_E = \max_{x \in E} F_n^A(x;E) = K_n^A . \qquad (2.2.7)$$

Definition 2.2.4

By a generalized lemniscatic region $\mathcal{L}_{F_n^A}$ we shall mean $\{x \in X \mid F_n^A(x;E) \le K_n^A\}$, and by $\hat{\mathcal{L}}_{F_n^A}$ we shall mean $\{x \in X \mid F_n^A(x;E) = K_n^A\}$.

(We do not notate this as $\partial\mathcal{L}_{F_n^A}$ as we did in the classical case since we do not necessarily have a maximum principle in the metric space $(X,\rho)$ ).

One can easily verify

Theorem 2.2.5

a) $E \subset \mathcal{L}_{F_n^A}$ and

b) $\hat{\mathcal{L}}_{F_n^A} \cap E \neq \emptyset$ .

Unlike the classical analogue, generalized Fekete points of order $n$ for $E$ do not have to lie on the boundary of $E$ . For example, in a discrete metric space $(X,\rho)$ and for a set $E$ of infinite cardinality, any set of $n$ distinct points of $E$ are generalized Fekete points of order $n$ for $E$ . It will sometimes happen that $\delta^A(E) = \delta^A(\partial E)$ , but this is no longer a certainty.

In the classical case, the Chebyshev constant could be viewed as the limiting value of the function $\tau_n(E)$ obtained by considering a min-max problem. It is through this approach that Hille extends the definition of Chebyshev constant for a compact set in an arbitrary metric space by means of a generalized average function.

Although Hille considers certain properties of generalized transfinite diameters, he completely neglects any discussion of generalized Chebyshev constants, beyong showing their existence and the result of Theorem 2.4.1. In Section 2.3 we shall discuss some of the consequences of the definition as well as define a generalized Chebyshev function by means of the generating function for the average.

## 2.3 A Generalized Chebyshev Constant.

Let $(X,\rho)$ be a metric space, $E$ a compact subset of $X$ , and

$A(u_1,u_2,\ldots,u_n)$ an average function. For $x_1,x_2,\ldots,x_n$ , $n \geq 1$ points of $X$ , and $x$ a variable point in $E$ , we let $\rho(x,x_i)$ denote the distances from $x$ to $x_i$ , $i = 1,2,\ldots,n$ . Let

$$\tau_n^A(E) = \min_{\substack{x_i \in X \\ i=1,2,\ldots,n}} \max_{x \in E} A(\rho(x,x_1),\rho(x,x_2),\ldots,\rho(x,x_n))^{1} \qquad (2.3.1)$$

Now $\lim\limits_{n\to\infty} \tau_n^A(E)$ exists, (see Hille [16]) and we have

Definition 2.3.1

$\chi^A(E) = \lim\limits_{n\to\infty} \tau_n^A(E)$ is called a generalized Chebyshev constant for $E$ .

A generalized Chebyshev constant for any average $A$ is monotone with respect to subsets, and if $A$ is a $\lambda$-th power average, then it is monotone with respect to $\lambda$ . Thus

Proposition 2.3.1

a) If $E_1 \subset E_2$ then $\chi^A(E_1) \leq \chi^A(E_2)$ .

b) If $\lambda < \mu$ then $\chi^{(\lambda)}(E) \leq \chi^{(\mu)}(E)$ .

The proof of Proposition 2.3.1 part a) follows easily from the definition of $\chi^A(E)$ , and the proof of part b) follows from Theorem 2.1.2 part a).

Definition 2.3.2

Points $\overset{*}{x}_1,\overset{*}{x}_2,\ldots,\overset{*}{x}_n$ such that

---

[1] If $E$ is merely bounded we define

$$\tau_n^A(E) = \inf_{\substack{x_i \in X \\ i=1,2,\ldots,n}} \sup_{x \in E} A(\rho(x,x_1),\rho(x,x_2),\ldots,\rho(x,x_n))$$

$$\min_{\substack{x_i \in X \\ i=1,2,\ldots,n}} \max_{x \in E} A(\rho(x,x_1),\rho(x,x_2),\ldots,\rho(x,x_n)) = \max_{x \in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*))$$

are called generalized Chebyshev points of order $n$ for $E$ .

Suppose $A$ is generated by $G$ , and $G(t)$ is increasing.

Definition 2.3.3

The function $C_n^A(x;E) = G^{-1}[nG(A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)))]$ is called a generalized Chebyshev function of order $n$ for $E$ .

Now $C_n^A(x;E) > 0$ . Let

$$\|C_n^A(x;E)\|_E = \max_{x \in E} C_n^A(x;E) = m_n^A(E) \tag{2.3.2}$$

Theorem 2.3.1

$m_n^A(E) = G^{-1}[nG(\tau_n^A(E))]$ .

We omit the proof as it is similar to the proof of Theorem 2.2.2. We note in the classical case this reduces to $(m_n(E))^{\frac{1}{n}} = \tau_n(E)$ .

Definition 2.3.4

We may again define a generalized lemniscatic region $\mathcal{L}_{C_n^A}$ by $\{x \in X \mid C_n^A(x;E) \le m_n^A(E)\}$ and a sub-region $\hat{\mathcal{L}}_{C_n^A}$ by $\{x \in X \mid C_n^A(x;E) = m_n^A(E)\}$ .

One can show:

Theorem 2.3.2

a) $E \subset \mathcal{L}_{C_n^A}$ and

b) $E \cap \hat{\mathcal{L}}_{C_n^A} \neq \emptyset$ .

The two most important theorems regarding Chebyshev points in the classical case, Theorem 1.2.2, and Theorem 1.2.3, no longer hold in the general case. For if $(X,\rho)$ is a discrete metric space and $E$ a subset of infinite cardinality, then for every set of points $x_1,x_2,\dots,x_n \in X$ , we have $\max_{x\in E} A(\rho(x,x_1),\rho(x,x_2),\dots,\rho(x,x_n)) = A(1,1,1,\dots,1) = 1$ . Hence $\chi^A(E) = 1$ . Thus a set of generalized Chebyshev points of order $n$ for $E$ is any set of $n$ points of $X$ ! Clearly they are not unique nor do they have to lie in the convex hull of $E$ .

However, we can obtain a "convex hull theorem" by putting additional structure on the metric space. To this end we note that the proof of Theorem 1.2.3 rested heavily on Fejér's Principle which does not hold in a general metric space. We do not have to use the pathological discrete metric space to illustrate this.

Example 2.3.1

Let $R_2^\infty$ be 2-dimensional real space with the metric given by: If $(x_1,y_1) \in R_2^\infty$ and $(x_2,y_2) \in R_2^\infty$ then $\rho((x_1,y_1),(x_2,y_2)) = \max(|x_1-x_2|,|y_1-y_2|)$ . Let $E = \{(x,y) \in R_2^\infty | \max(|x|,|y|) \le 1\}$ , i.e., the unit ball for $R_2^\infty$ . Let $Q_1$ be the point $(0,2)$ . We claim that for every point $P \in R_2^\infty$ , $\exists\, Q \in E \ni \rho(P,Q) \ge \rho(Q,Q_1)$ .

Let $S = \{(x,y) \in R_2^\infty | x \le 0\}$ and let $S' = \{(x,y) \in R_2^\infty | x > 0\}$ . Thus $S \cup S' = R_2^\infty$ . For every $P \in S$ , let $Q = (1,1)$ . Then $\rho(P,Q) = \max(|x-1|,|y-1|) \ge 1 = \max(|0-1|,|2-1|) = \rho(Q,Q_1)$ . For every $P \in S'$ , let $Q = (-1,1)$ . Then

$\rho(P,Q) = \max(|x+1|, |y-1|) > 1 = \max(|0+1|, |2-1|) = \rho(Q,Q_1)$ .

Shisha [27] shows that Fejér's Principle holds in an n-dimensional real Euclidean space. We can show more:

Definition 2.3.5

Let $(X,\rho)$ be a metric space. A subset $V$ of $X$ is called proximinal if and only if every element $x \in X$ has a best approximation out of $V$ , i.e., for every $x \in X$ , there exists a $\bar{v} \in V$ such that $\rho(x,\bar{v}) = \inf_{v \in V} \rho(x,v)$ .

Theorem 2.3.3

Fejér's Principle is satisfied in inner-product spaces for subsets which are both proximinal and convex.

Proof: Let $Y$ be an inner-product space and let $V$ be a proximinal and convex subset of $Y$ . Let $x^*$ be in $Y \backslash V$ . We shall show there exists a $\bar{v} \in V$ such that $\|v-\bar{v}\| < \|v-x^*\|$ for every $v \in V$ .

Define $\bar{v}$ to be a best approximation to $x^*$ out of $V$ . Since $V$ is proximinal, $\bar{v}$ exists. The following characterization is given in Cheney and Goldstein [2]:

$\bar{v}$ is a best approximation to $x^*$ out of $V$ if and only if $(x^*-\bar{v},\bar{v}-v) \geq 0$ for every $v$ in a closed convex subset $V$ . Now $\|x^*-v\|^2 = \|(x^*-\bar{v}) + (\bar{v}-v)\|^2 = \|x^*-\bar{v}\|^2 + \|\bar{v}-v\|^2 + 2(x^*-\bar{v},\bar{v}-v)$ . Since the last term on the right is non-negative we have $\|x^*-v\|^2 > \|\bar{v}-v\|^2$ or $\|x^*-v\| > \|\bar{v}-v\|$ .

If $Y$ is a metric space, then any compact set $E$ is proximinal. If $H$ is a Hilbert space, then any closed convex set is proximinal. Proximinal subsets of other spaces may be found in the Appendices of Singer [28].

### Theorem 2.3.4

Let $(X,\rho)$ be a Hilbert space where the metric is derived from the inner product. Let $E$ be a compact subset of $X$ . Then every set of generalized Chebyshev points of $E$ lies in $V = K(E)$ , the closed convex hull of $E$ .

Proof: Let $x_i^*$, $i = 1,2,\ldots,n$ be a set of generalized Chebyshev points for $E$ with respect to an average $A$ . Thus

$$\max_{X\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)) \le \max_{X\in E} A(\rho(x,x_1),\rho(x,x_2),\ldots,\rho(x,x_n)) \tag{2.3.3}$$

for any set $x_1,x_2,\ldots,x_n \in X$ . Suppose $x_1^* \in X\backslash V$ . Let $\inf_{v\in V} \rho(x_1^*,v) = \rho(x_1^*,\bar{v})$ . Such a $\bar{v}$ exists since $V$ is a proximinal subset of $X$ . Since $V$ is also convex, the preceding theorem (with the substitution $x$ for $v$ and $x_1^*$ for $x^*$) gives us $\rho(x,\bar{v}) < \rho(x,x_1^*)$ for every $x \in V$ and thus for every $x \in E$ . POSTULATE b) gives $A(\rho(x,\bar{v}),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)) < A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*))$ for every $x \in E$ . Thus

$$\max_{X\in E} A(\rho(x,\bar{v}),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)) < \max_{X\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*))$$

contrary to (2.3.3).

### Corollary 2.3.1

If $E$ is a convex compact subset of a Hilbert space, then every set of generalized Chebyshev points of $E$ lie in $E$ .

The fact that a set of Chebyshev points of $E$ lies in $E$ does not imply that $E$ is convex. Consider:

**Example 2.3.2**

Let $E = \{\partial D \cup 0\}$ where $0$ is the center of $D$. In the classical case the Chebyshev points of $E$ are $x_1^* = x_2^* = \ldots = x_n^* = 0$ for all $n$. Thus they lie in $E$, but $E$ is not convex.

In Chapter 4 we shall say more about Fejér's Principle and the Convex Hull Theorem.

In the general case we do not have Fekete's Theorem 1.3.1. We have a much weaker result.

## 2.4 The Basic Inequality

**Theorem 2.4.1**

For any compact set $E$ in a metric space $(X,\rho)$, and for any average $A(x_1,x_2,\ldots,x_n)$, we have $\chi^A(E) \le \delta^A(E)$.

Proof: Let $G(t)$ be an increasing generating function for $A(x_1,x_2,\ldots,x_n)$. Let $\xi_1,\xi_2,\ldots,\xi_n$ be a set of Fekete points of order $n$ for $E$, and let $x_1^*,x_2^*,\ldots,x_n^*$ be a set of Chebyshev points of order $n$ for $E$. Now

$$\max_{x\in E} A_G(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)) \le \max_{x\in E} A_G(\rho(x,\xi_1),\rho(x,\xi_2),\ldots,\rho(x,\xi_n))$$

$$= A_G(\rho(\bar{x},\xi_1),\rho(\bar{x},\xi_2),\ldots,\rho(\bar{x},\xi_n)) \qquad (2.4.1)$$

$$V_{n+1}^A(E) = \max_{\substack{x_i\in E\\ i=1,2,\ldots,n+1}} V(x_1,x_2,\ldots,x_{n+1}) \ge V^A(\xi_1,\xi_2,\ldots,\xi_n,\bar{x})$$

$$= G^{-1}\left[\binom{n+1}{2}G(A_G(\rho(\xi_i,\xi_j),\rho(\bar{x},\xi_k)))\right]$$

for $1 \le i < j \le n$, and $k = 1,2,\ldots,n$. Replacing the arguments by their averages, we have:

$V^A_{n+1}(E) \geq G^{-1}\left[\binom{n+1}{2}G(A(A_G(\rho(\xi_i,\xi_j)),\ A_G(\rho(\bar{x},\xi_k))))\right]$ . From (2.2.4), (2.4.1), (2.3.1), and POSTULATE b) this becomes $V^A_{n+1}(E) \geq G^{-1}\left[\binom{n+1}{2}G(A_G(d^A_n(E)\ ,\tau^A_n(E)))\right]$ or $d^A_{n+1}(E) \geq A_G(d^A_n(E),\tau^A_n(E)) \geq \min\ (d^A_n(E),\tau^A_n(E))$ . Either $d^A_n(E) = \tau^A_n(E)$ which implies $\delta^A(E) = \chi^A(E)$ or from (2.1.3), $d^A_{n+1}(E) > \min\ (d^A_n(E),\tau^A_n(E))$ . If $d^A_n(E) < \tau^A_n(E)$ , then $d^A_{n+1}(E) > d^A_n(E)$ contrary to Theorem 2.2.3. Hence $\tau^A_n(E) \leq d^A_n(E)$ and $\chi^A(E) \leq \delta^A(E)$ .

## 2.5 Additional Observations

The definitions of restricted Chebyshev constant, restricted Chebyshev points, and restricted Chebyshev polynomial may also be extended to an arbitrary metric space $(X,\rho)$ . Let $E$ be a compact subset of $X$ , and $A(x_1,x_2,\ldots,x_n)$ an average. Let

$$\tilde{\tau}^A_n(E) = \min_{\substack{x_i \in E \\ i=1,2,\ldots,n}} \max_{x \in E} A(\rho(x,x_1),\rho(x,x_2),\ldots,\rho(x,x_n)) \tag{2.5.1}$$

The limit of $\tilde{\tau}^A_n(E)$ exists as $n \to \infty$ and we have

**Definition 2.5.1**

$\tilde{\chi}^A(E) = \lim\limits_{n\to\infty} \tilde{\tau}^A_n(E)$ is called a generalized restricted Chebyshev constant of order $n$ for $E$ .

**Definition 2.5.2**

Points $\tilde{x}^*_1,\tilde{x}^*_2,\ldots,\tilde{x}^*_n$ such that $\min\limits_{\substack{x_i \in E \\ i=1,2,\ldots,n}} \max\limits_{x \in E} A(\rho(x,x_1),\rho(x,x_2),\ldots,\rho(x,x_n))$

$= \max A(\rho(x,\tilde{x}_1^*), \rho(x,\tilde{x}_2^*), \ldots, \rho(x,\tilde{x}_n^*))$ are called *generalized restricted Chebyshev points of order $n$ for $E$*.

Suppose $A(x_1, x_2, \ldots, x_n)$ is generated by $G \in T$ and $G$ is increasing.

*Definition 2.5.3*

The function $\tilde{C}_n^A(x;E) = G^{-1}\left[nG(A_G(\rho(x,\tilde{x}_1^*), \rho(x,\tilde{x}_2^*), \ldots, \rho(x,\tilde{x}_n^*)))\right]$

is called a *generalized restricted Chebyshev function of order $n$ for $E$*.

Let

$$\|\tilde{C}_n^A(x;E)\|_E = \max_{x \in E} \tilde{C}_n^A(x) = \tilde{m}_n^A(E) \qquad (2.5.2)$$

*Theorem 2.5.1*

$$\tilde{m}_n^A(E) = G^{-1}\left[nG(\tilde{\tau}_n^A(E))\right]$$

In the generalized case we no longer have the equality of the Chebyshev constant and the restricted Chebyshev constant. From the definition, the following inequality results:

$$\chi^A(E) \leq \tilde{\chi}^A(E) \qquad (2.5.3)$$

In Chapter 3, Example 3.1.1, we shall see a case where strict inequality holds.

A theorem analagous to Theorem 1.2.3 is not applicable here, since by definition the generalized restricted Chebyshev points lie in $E$. However, we can make a few statements regarding the sets of generalized restricted Chebyshev points and the sets of generalized Chebyshev points for a compact set $E$ in $(X,\rho)$ .

### Theorem 2.5.2

If $x_1^*, x_2^*, \ldots, x_n^* \in E$ are generalized Chebyshev points of order $n$ for $E$ , then $x_1^*, x_2^*, \ldots, x_n^*$ are generalized restricted Chebyshev points of order $n$ for $E$ .

Proof:

$$\begin{aligned} &\max_{x\in E} A(\rho(x,x_1^*), \rho(x,x_2^*), \ldots, \rho(x,x_n^*)) \\ \leq\ &\max_{x\in E} A(\rho(x,x_1), \rho(x,x_2), \ldots, \rho(x,x_n)) \end{aligned} \tag{2.5.4}$$

for all $x_1, x_2, \ldots, x_n \in X$ . Since $E \subset X$ , (2.5.4) holds for all $x_1, x_2, \ldots, x_n \in E$ . Hence $x_1^*, x_2^*, \ldots, x_n^*$ are generalized restricted Chebyshev points of order $n$ for $E$ .

### Theorem 2.5.3

If $\tilde{x}_1^*, \tilde{x}_2^*, \ldots, \tilde{x}_n^* \in E$ are generalized restricted Chebyshev points of order $n$ for $E$ , and if there exists a set of generalized Chebyshev points of order $n$ for $E$ lying in $E$ , then $\tilde{x}_1^*, \tilde{x}_2^*, \ldots, \tilde{x}_n^*$ are generalized Chebyshev points of order $n$ for $E$ .

Proof:

$$\max_{x\in E} A(\rho(x,\tilde{x}_1^*), \rho(x,\tilde{x}_2^*), \ldots, \rho(x,\tilde{x}_n^*)) \leq \max_{x\in E} A(\rho(x,x_1), \rho(x,x_2), \ldots, \rho(x,x_n))$$

for all $x_1, x_2, \ldots, x_n \in E$ . If $x_1^*, x_2^*, \ldots, x_n^* \in E$ are a set of generalized Chebyshev points of order $n$ for $E$ , then

$$\max_{x\in E} A(\rho(x,\tilde{x}_1^*),\rho(x,\tilde{x}_2^*),\ldots,\rho(x,\tilde{x}_n^*)) \le \max_{x\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)) \tag{2.5.5}$$

But

$$\max_{x\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)) \le \max_{x\in E} A(\rho(x,x_1),\rho(x,x_2),\ldots,\rho(x,x_n))$$

for all $x_1,x_2,\ldots,x_n \in X$ . Hence

$$\max_{x\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)) \le \max_{x\in E} A(\rho(x,\tilde{x}_1^*),\rho(x,\tilde{x}_2^*),\ldots,\rho(x,\tilde{x}_n^*)) \tag{2.5.6}$$

From (2.5.5) and (2.5.6) we get

$$\max_{x\in E} A(\rho(x,\tilde{x}_1^*),\rho(x,\tilde{x}_2^*),\ldots,\rho(x,\tilde{x}_n^*)) = \max_{x\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*))$$

and $\tilde{x}_1^*,\tilde{x}_2^*,\ldots,\tilde{x}_n^*$ are generalized Chebyshev points of order $n$ for $E$ .

Corollary 2.5.1

If $E$ is a convex compact subset of a Hilbert space, then for every $n$ , every set of generalized Chebyshev points of order $n$ for $E$ is a set of generalized restricted Chebyshev points of order $n$ for $E$ , and conversely.

Corollary 2.5.1 indicates that in the case of convex compact subsets of Hilbert spaces, the unicity of generalized Chebyshev points for each $n$ implies the unicity of generalized restricted Chebyshev points, and conversely.

If the set $E$ is not convex, then the uniqueness question for the sets of generalized Chebyshev points of order $n$ for $E$ is independent of the uniqueness question for the sets of generalized restricted Chebyshev points of order $n$ for $E$. To illustrate this, consider Example 1.5.1. Here, for each $n$, the Chebyshev points are unique, but the restricted Chebyshev points are not unique. In Example 3.1.1 of Chapter 3 we will illustrate a case where the generalized restricted Chebyshev points of order $n$ ($n$ even) are unique, whereas the generalized Chebyshev points of order $n$ ($n$ even) are not unique.

Besides Corollary 2.3.1 we have the following theorem which states sufficient conditions for generalized Chebyshev points of order $n$ for $E$ to lie in $E$.

Theorem 2.5.4

If $\tilde{m}_n^A(E) = m_n^A(E)$ for every $n$ and the generalized Chebyshev points of order $n$ for $E$ are unique, then they lie in $E$.

Proof: Let $x_1^*, x_2^*, \dots, x_n^*$ be the set of generalized Chebyshev points of order $n$ for $E$. Their uniqueness implies

$$\begin{aligned} &\max_{x\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\dots,\rho(x,x_n^*)) \\ \le\ &\max_{x\in E} A(\rho(x,x_1),\rho(x,x_2),\dots,\rho(x,x_n)) \end{aligned} \tag{2.5.7}$$

for every set $x_1,x_2,\dots,x_n \in X$ with equality if and only if $\{x_1,x_2,\dots,x_n\} = \{x_1^*,x_2^*,\dots,x_n^*\}$. Since $\tilde{m}_n^A(E) = m_n^A(E)$ for every $n$, if $\tilde{x}_1^*,\tilde{x}_2^*,\dots,\tilde{x}_n^*$ are a set of generalized restricted Chebyshev points of order $n$ for $E$, we have

$$\begin{aligned} &\max_{x\in E} A(\rho(x,\tilde{x}_1^*),\rho(x,\tilde{x}_2^*),\ldots,\rho(x,\tilde{x}_n^*)) \\ = &\max_{x\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*)) \end{aligned} \qquad (2.5.8)$$

Since (2.5.7) holds for all sets $x_1,x_2,\ldots,x_n \in X$ , it must hold for all sets $x_1,x_2,\ldots,x_n \in E$ . Thus for $\tilde{x}_1^*,\tilde{x}_2^*,\ldots,\tilde{x}_n^* \in E$ , (2.5.8) and (2.5.7) give

$\max_{x\in E} A(\rho(x,\tilde{x}_1^*),\rho(x,\tilde{x}_2^*),\ldots,\rho(x,\tilde{x}_n^*)) = \max_{x\in E} A(\rho(x,x_1^*),\rho(x,x_2^*),\ldots,\rho(x,x_n^*))$ $\le \max A(\rho(x,\tilde{x}_1^*),\rho(x,\tilde{x}_2^*),\ldots,\rho(x,\tilde{x}_n^*))$ . Hence equality throughout implies $\{x_1^*,x_2^*,\ldots,x_n^*\} = \{\tilde{x}_1^*,\tilde{x}_2^*,\ldots,\tilde{x}_n^*\}$ and $x_1^*,x_2^*,\ldots,x_n^* \in E$ .

## CHAPTER 3

## Euclidean Spaces

Let $R_n$ be the set of real n-tuples. If $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ we make $R_n$ into a vector space by defining $x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$ and $\alpha x = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$ for any real number $\alpha$. An inner product structure is imposed on $R_n$ by defining $(x,y) = \sum_{i=1}^{n} x_i y_i$. With these definitions, $R_n$ is called real Euclidean n-space.

A glance at the Appendix shows that Pólya and Szegö [25] considered certain compact sets in $R_1$, $R_2$, and $R_3$, and calculated their transfinite diameters and Chebyshev constants for $\lambda$-th power averages. We shall restrict ourselves to averages for which $\lambda \geq 1$, but will obtain some results for all compact sets in $R_n$. Specifically we will calculate the Chebyshev constant for any compact set E in $R_n$. This will enable us to compare it with the transfinite diameter of E in certain known cases. We may then comment when we have equality, or strict inequality, between the two set functions. This question provided one motivation for this dissertation.

Since we are dealing with $\lambda$-th power averages, $\lambda \geq 1$, the Chebyshev constant and the transfinite diameter discussed in Chapter 3 should be called "generalized" in accordance with the definitions in Chapter 2. However, to avoid excess verbiage, we shall omit this adjective.

We start with

### 3.1 $R_1$ - The Real Line

Throughout this section E will denote a compact set in $R_1$.

We let

$$a = \min_{x \in E} x \qquad \text{and} \qquad b = \max_{x \in E} x \qquad (3.1.1)$$

We first prove the following lemma which will be quite useful in the proofs of this section.

Lemma 3.1.1

For every $x \in E$ and for $\lambda > 1$ we have $(b-a)^{\lambda} > (b-x)^{\lambda} + (x-a)^{\lambda}$ unless $x = a$ or $x = b$ in which case we have equality.

Proof: It is well known that the p-norm is a strictly decreasing function of p for $p > 1$ (See Hardy, Littlewood, and Pólya [11]).

Thus for $x_i \geq 0$, $\sum_{i=1}^{n} x_i > \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}}$ unless all but one of the $x_i$'s are zero. Let $(b-x) = x_1$ and $(x-a) = x_2$. For $n = 2$ and $p = \lambda$, the conclusion follows.

We note that for $\lambda = 1$, we have equality.

Theorem 3.1.1

For $\lambda \geq 1$, and n even, the Fekete points of order n for E are $\{\underbrace{a,a,\ldots,a}_{\frac{n}{2}}\,;\,\underbrace{b,b,\ldots,b}_{\frac{n}{2}}\}$ and they are unique.

Proof: We shall prove by induction $d_n^{(\lambda)}(E) \leq \left[\frac{1}{\binom{n}{2}} \frac{n^2}{4} (b-a)^{\lambda}\right]^{\frac{1}{\lambda}}$.

For $n = 2$, $d_2^{(\lambda)}(E) = \max |x_i - x_j| = (b-a) \leq (b-a)$. (3.1.2)

Suppose it is true for $n = k$ , k even. Hence

$$d_k^{(\lambda)}(E) = \max \left[\frac{1}{\binom{k}{2}} \sum_{1 \le i < j \le k} |x_i - x_j|^{\lambda}\right]^{\frac{1}{\lambda}} \le \left[\frac{1}{\binom{k}{2}} \frac{k^2}{4} (b-a)^{\lambda}\right]^{\frac{1}{\lambda}} \tag{3.1.3}$$

Consider $k + 2$ points. Writing $x_1 \ge x_2 \ge \ldots \ge x_{k+2}$ , we can dispense with absolute values. Now

$$\left[d_{k+2}^{(\lambda)}(E)\right]^{\lambda} = \max \left[\frac{1}{\binom{k+2}{2}} \sum_{1 \le i < j \le k+2} (x_i - x_j)^{\lambda}\right].$$

We break up the sum so as to consider those terms involving $x_1$ and $x_{k+2}$ separately.

Thus

$$\left[d_{k+2}^{(\lambda)}(E)\right]^{\lambda} = \max \frac{1}{\binom{k+2}{2}} \left[\sum_{2 \le i < j \le k+1} (x_i - x_j)^{\lambda} + \sum_{i=2}^{k+1} (x_1 - x_i)^{\lambda} + \sum_{i=1}^{k+1} (x_i - x_{k+2})^{\lambda}\right]$$

$$= \max \frac{1}{\binom{k+2}{2}} \left[\sum_{2 \le i < j \le k+1} (x_i - x_j)^{\lambda} + \sum_{i=2}^{k+1} (x_1 - x_i)^{\lambda} + (x_i - x_{k+2})^{\lambda} + (x_1 - x_{k+2})^{\lambda}\right]$$

By Lemma 3.1.1 we have

$$(x_1 - x_i)^{\lambda} + (x_i - x_{k+2})^{\lambda} \le (x_1 - x_{k+2})^{\lambda} \tag{3.1.4}$$

for $i = 2, 3, \ldots, k+1$ . Thus

$$\left[d_{k+2}^{(\lambda)}(E)\right]^{\lambda} \le \max \frac{1}{\binom{k+2}{2}} \left[\sum_{2\le i<j\le k+1} (x_i - x_j)^{\lambda} + (k+1)(x_1 - x_{k+2})^{\lambda}\right]$$

$$\le \frac{1}{\binom{k+2}{2}} \left[\max \sum_{2\le i<j\le k+1} (x_i - x_j)^{\lambda} + \max\ (k+1)(x_1 - x_{k+2})^{\lambda}\right].$$

By (3.1.3) the first term which involves $k$ points is less than or equal to $\frac{k^2}{4}(b-a)^{\lambda}$. By (3.1.2) the second term involving $2$ points is equal to $(k+1)(b-a)^{\lambda}$. Thus

$$\left[d_{k+2}^{(\lambda)}(E)\right]^{\lambda} \le \frac{1}{\binom{k+2}{2}} \left[\frac{k^2}{4}(b-a)^{\lambda} + (k+1)(b-a)^{\lambda}\right] = \frac{1}{\binom{k+2}{2}} \frac{(k+2)^2}{4} (b-a)^{\lambda}$$

or

$$d_{k+2}^{(\lambda)}(E) \le \left[\frac{1}{\binom{k+2}{2}} \frac{(k+2)^2}{4} (b-a)^{\lambda}\right]^{\frac{1}{\lambda}} \tag{3.1.5}$$

Thus by induction, for every $n$,

$$d_n^{(\lambda)}(E) \le \left[\frac{1}{\binom{n}{2}} \frac{n^2}{4} (b-a)^{\lambda}\right]^{\frac{1}{\lambda}} \tag{3.1.6}$$

But if we choose $x_i = b$ for $i = 1,2,\ldots,\frac{n}{2}$ and $x_i = a$ for $i = \frac{n}{2}+1,\ldots,n$ then

$$d_n^{(\lambda)}(E) = \max\left[\frac{1}{\binom{n}{2}} \sum_{1\le i<j\le n} (x_i - x_j)^{\lambda}\right]^{\frac{1}{\lambda}} \ge \left[\frac{1}{\binom{n}{2}} \frac{n^2}{4} (b-a)^{\lambda}\right]^{\frac{1}{\lambda}}. \tag{3.1.7}$$

Hence, (3.1.6) and (3.1.7) give

$$d_n^{(\lambda)}(E) = \left[\frac{1}{\binom{n}{2}} \frac{n^2}{4} (b-a)^{\lambda}\right]^{\frac{1}{\lambda}} \tag{3.1.8}$$

and $\{\underbrace{a,a,\dots,a}_{\frac{n}{2}}\;;\;\underbrace{b,b,\dots,b}_{\frac{n}{2}}\}$ is a set of Fekete points of order $n$, $n$ even for $E$. The uniqueness follows from the fact that in (3.1.4) we have strict inequality unless $x_i = x_1$ or $x_{k+2}$ for $i = 2,3,\dots,k+1$, in the case $\lambda > 1$. In the case $\lambda = 1$ uniqueness comes from the strict inequality $(x_1 - x_{k+2}) < (b - a)$ unless $x_1 = b$ and $x_{k+2} = a$.

Corollary 3.1.1

For $\lambda \geq 1$, $\delta^{(\lambda)}(E) = \dfrac{(b-a)}{2^{\frac{1}{\lambda}}}$

Proof: Result follows from (3.1.8) as $n \to \infty$.

Theorem 3.1.2

For $n$ odd, Fekete points of $E$ are not unique. For $\lambda = 1$, any set of the form $S = \{\underbrace{a,a,\dots,a}_{\frac{n-1}{2}};\bar{\xi};\underbrace{b,b,\dots,b}_{\frac{n-1}{2}}\}$ where $\bar{\xi} \in E$ is a set of Fekete points of order $n$ for $E$. For $\lambda > 1$, for each odd $n$ there exist two sets of Fekete points of order $n$ for $E$, namely set $S$ where $\bar{\xi} = a$ and set $S$ where $\bar{\xi} = b$.

Proof: Again let $x_1 \geq x_2 \geq \dots \geq x_n$, and $\bar{i} = \frac{n+1}{2}$. We shall denote $x_{\bar{i}}$ by $\bar{\xi}$. Then

$$\left[d_n^{(\lambda)}(E)\right] = \frac{1}{\binom{n}{2}} \max \left[\sum_{1\leq i<j\leq n} (x_i - x_j)^{\lambda}\right]$$

$$= \frac{1}{\binom{n}{2}} \max \left[\sum_{1\leq i<j\leq n} (x_i - x_j)^{\lambda} + \sum_{i=1}^{\frac{n-1}{2}} (x_i - \bar{\xi})^{\lambda} + \sum_{j=\frac{n+3}{2}}^{n} (\bar{\xi} - x_j)^{\lambda}\right]$$

From Lemma 3.1.1 and Theorem 3.1.1 we have

$$\left[d_n^{(\lambda)}(E)\right]^{\lambda} \le \frac{1}{\binom{n}{2}}\left[\frac{(n-1)^2}{4}(b-a)^{\lambda} + \frac{(n-1)}{2}(b-a)^{\lambda}\right]$$

or

$$d_n^{(\lambda)}(E) \le \left[\frac{1}{\binom{n}{2}}\left(\frac{(n-1)^2}{4} + \frac{n-1}{2}\right)(b-a)^{\lambda}\right]^{\frac{1}{\lambda}}.$$ But if we choose

$x_i = b$ for $i = 1,2,\ldots,\frac{n-1}{2}$, and $x_i = a$ for $i = \frac{n+1}{2},\ldots,n$ or

$x_i = b$ for $i = 1,2,\ldots,\frac{n+1}{2}$ and $x_i = a$ for $i = \frac{n+3}{2},\ldots,n$ we have

$$d_n^{(\lambda)}(E) \ge \left[\frac{1}{\binom{n}{2}}\left(\frac{n-1}{2}\right)\left[\frac{n-1}{2}+1\right](b-a)^{\lambda}\right]^{\frac{1}{\lambda}}.$$

Hence for $\lambda > 1$ the Fekete points of order $n$, $n$ odd, are as described. These are the only sets since the condition for equality in the lemma again rules out any other sets. In the case $\lambda = 1, (x_i-\bar{\xi}) + (\bar{\xi}-x_j) = (x_i-x_j)$, so $\bar{\xi}$ can be any point of $E$.

From Definition 2.2.3 a generalized Fekete function of order $n$ for $E$ takes the form $F_n^{(\lambda)}(x;E) = \left[\sum_{i=1}^{n} |x-\xi_i|^{\lambda}\right]^{\frac{1}{\lambda}}$ where $\xi_i$ $i = 1,2,\ldots,n$ are generalized Fekete points of order $n$ for $E$.

Let $n$ be even. For $\lambda \ge 1$, the Fekete function of order $n$ for $E$ is unique. It is

$$F_n^{(\lambda)}(x;E) = \left(\frac{n}{2}\right)^{\frac{1}{\lambda}}\left[|x-a|^{\lambda} + |x-b|^{\lambda}\right]^{\frac{1}{\lambda}} \tag{3.1.9}$$

From Definition 2.2.4 we see that in this case $K_n^{(\lambda)} = \|F_n^{(\lambda)}(x;E)\|_E =$ $\left(\frac{n}{2}\right)^{\frac{1}{\lambda}}(b-a)$ and

$$\mathcal{L}_{F_n^{(\lambda)}} : \left\{x \in R_1 \,\middle|\, \left(\frac{n}{2}\right)^{\frac{1}{\lambda}} \left[\,|x-a|^\lambda + |x-b|^\lambda\right]^{\frac{1}{\lambda}} \le \left(\frac{n}{2}\right)^{\frac{1}{\lambda}} (b-a)\right\} = [a,b] \qquad (3.1.10)$$

For the case $\lambda = 1$ we have

$$\hat{\mathcal{L}}_{F_n^{(1)}} = \mathcal{L}_{F_n^{(1)}} = [a,b] \qquad (3.1.11)$$

and for $\lambda > 1$ we have

$$\hat{\mathcal{L}}_{F_n^{(\lambda)}} = \{a,b\}\,. \qquad (3.1.12)$$

Let $n$ be odd. For $\lambda > 1$ and for each $n$, there exist two Fekete functions of order $n$ for $E$. Corresponding to $\{\underbrace{a,a,\ldots,a}_{\frac{n+1}{2}}\,;\,\underbrace{b,b,\ldots,b}_{\frac{n-1}{2}}\}$ we have

$$F_n^{(\lambda)}(x;E) = \left[\left(\frac{n+1}{2}\right)|x-a|^\lambda + \left(\frac{n-1}{2}\right)|x-b|^\lambda\right]^{\frac{1}{\lambda}} \qquad (3.1.13)$$

and corresponding to $\{\underbrace{a,a,\ldots,a}_{\frac{n-1}{2}}\,;\,\underbrace{b,b,\ldots,b}_{\frac{n+1}{2}}\}$ we have

$$F_n^{(\lambda)}(x;E) = \left[\left(\frac{n-1}{2}\right)|x-a|^\lambda + \left(\frac{n+1}{2}\right)|x-b|^\lambda\right]^{\frac{1}{\lambda}} \qquad (3.1.14)$$

In both cases $K_n^{(\lambda)} = \left(\frac{n+1}{2}\right)^{\frac{1}{\lambda}}(b-a)$ and

$$\mathcal{L}_{F_n^{(\lambda)}} = [a-\epsilon(n), b] \qquad (3.1.15)$$

and

$$\hat{\mathcal{L}}_{F_n^{(\lambda)}} = \{a-\epsilon(n), b\} \qquad (3.1.16)$$

for $F_n^{(\lambda)}(x;E)$ given by (3.1.13), while

$$\mathcal{L}_{F_n^{(\lambda)}} = [a, b+\epsilon(n)] \qquad (3.1.17)$$

and

$$\hat{\mathcal{L}}_{F_n^{(\lambda)}} = \{a, b+\epsilon(n)\} \qquad (3.1.18)$$

for $F_n^{(\lambda)}(x;E)$ given by (3.1.14).

For $n$ odd and $\lambda = 1$, Fekete points of order $n$ for $E$ are $\{\underbrace{a,a,\dots,a}_{\frac{n-1}{2}};\ \bar{\xi}\ ;\ \underbrace{b,b,\dots,b}_{\frac{n-1}{2}}\}$ where $\bar{\xi}$ is any point in $E$. For each $\bar{\xi} \in E$, there exists a Fekete function of order $n$ for $E$ as follows:

$$F_n^{(1)}(x;E) = \left[\left(\frac{n-1}{2}\right)|x-a| + |x-\xi| + \left(\frac{n-1}{2}\right)|x-b|\right] \qquad (3.1.19)$$

Let $\bar{\zeta} = \max\left(|b-\bar{\xi}|, |a-\bar{\xi}|\right)$. We see $\bar{\zeta} \le |b-a|$ for every $\bar{\xi} \in E$. Hence $K_n^{(\lambda)} = \left(\frac{n-1}{2}\right)(b-a) + \bar{\zeta}$. We have

$$\mathcal{L}_{F_n}(1) = \begin{cases} \left[a, b + \dfrac{2\bar{\xi}-(a+b)}{n}\right] & \text{for} \quad \dfrac{a+b}{2} \le \bar{\xi} \le b \\ \left[a - \dfrac{(b+a-2\bar{\xi})}{n}, b\right] & \text{for} \quad a \le \bar{\xi} \le \dfrac{a+b}{2} \\ \left[a, b\right] \quad \text{for} \quad \bar{\xi} = \dfrac{a+b}{2} & \end{cases} \tag{3.1.20}$$

and

$$\hat{\mathcal{L}}_{F_n}(1) = \begin{cases} \left\{a, b + \dfrac{2\bar{\xi}-(a+b)}{n}\right\} & \text{for} \quad \dfrac{a+b}{2} < \bar{\xi} \le b \\ \left\{a - \dfrac{(b+a-2\bar{\xi})}{n}, b\right\} & \text{for} \quad a \le \bar{\xi} \le \dfrac{a+b}{2} \\ \left\{a, b\right\} \quad \text{for} \quad \bar{\xi} = \dfrac{a+b}{2} & \end{cases} \tag{3.1.21}$$

We now turn our attention to the problem of finding the Chebyshev constant for any compact set $E$ in $R_1$ . We have the following result:

<u>Theorem 3.1.3</u>

For $\lambda \ge 1$ , $\chi^{(\lambda)}(E) = \frac{1}{2}(b-a)$ .

<u>**Proof**</u>: $\tau_n^{(\lambda)}(E) = \min\limits_{\substack{x_i \in R_1 \\ i=1,2,\ldots,n}} \max\limits_{x \in E} \left[\frac{1}{n} \sum\limits_{i=1}^{n} |x-x_i|^{\lambda}\right]^{\frac{1}{\lambda}}$ . Put $x_i = \frac{b+a}{2}$ for $i = 1,2,\ldots,n$ . Then $\tau_n^{(\lambda)}(E) \le \max\limits_{x \in E} \left[\frac{1}{n} n \left|x - \frac{(b+a)}{2}\right|^{\lambda}\right]^{\frac{1}{\lambda}}$ or

$\tau_n^{(\lambda)}(E) \le \frac{1}{2}(b-a)$ .

Now for $\lambda \ge 1$ , Minkowski's inequality gives

$$n^{\frac{1}{\lambda}}(b-a) = \left(\sum_{i=1}^{n} |b-x_i+x_i-a|^{\lambda}\right)^{\frac{1}{\lambda}} \le \left(\sum_{i=1}^{n} (|b-x_i| + |x_i-a|)^{\lambda}\right)^{\frac{1}{\lambda}}$$

$$\le \left(\sum_{i=1}^{n} |b-x_i|^{\lambda}\right)^{\frac{1}{\lambda}} + \left(\sum_{i=1}^{n} |x_i-a|^{\lambda}\right)^{\frac{1}{\lambda}} .$$

Thus either

$$\left[\frac{1}{n}\sum_{i=1}^{n}|b-x_i|^{\lambda}\right]^{\frac{1}{\lambda}} \geq \tfrac{1}{2}(b-a) \quad \text{or} \quad \left[\frac{1}{n}\sum_{i=1}^{n}|a-x_i|^{\lambda}\right]^{\frac{1}{\lambda}} \geq \tfrac{1}{2}(b-a) \ .$$

This is true for all $x_i \in R_1$ . Hence, if $x_i^*$, $i = 1,2,\ldots,n$ are Chebyshev points of order $n$ for $E$ , then we have

$$T_n^{(\lambda)}(E) = \max_{x\in E}\left[\frac{1}{n}\sum_{i=1}^{n}|x-x_i^*|^{\lambda}\right]^{\frac{1}{\lambda}} \geq \max_{x\in\{a,b\}}\left[\frac{1}{n}\sum_{i=1}^{n}|x-x_i^*|^{\lambda}\right]^{\frac{1}{\lambda}} \geq \tfrac{1}{2}(b-a) \ .$$

Hence for every $n$ , $T_n^{(\lambda)}(E) = \frac{1}{2}(b-a)$ and $\chi^{(\lambda)}(E) = \frac{1}{2}(b-a)$ .

## Theorem 3.1.4

For $\lambda > 1$ , $x_1^* = x_2^* = \ldots = x_n^* = \frac{b+a}{2}$ are Chebyshev points of order $n$ for $E$ and they are unique.

Proof: The proof of Theorem 3.1.3 shows that $x_i^* = \frac{b+a}{2}$ for $i = 1,2,\ldots,n$ are a set of Chebyshev points of order $n$ for $E$ . Their uniqueness follows from the fact that for $x_i^* \neq \frac{b+a}{2}$ for every $i$ , we have strict inequality in Minkowski's inequality. Thus, for either $x = a$ or $x = b$ we have $\left[\frac{1}{n}\sum_{i=1}^{n}|x-x_i|^{\lambda}\right]^{\frac{1}{\lambda}} > \frac{1}{2}(b-a)$ . Hence for $x_i \neq \frac{b+a}{2}$ , $\max_{x\in E}\left[\frac{1}{n}\sum_{i=1}^{n}|x-x_i|^{\lambda}\right]^{\frac{1}{\lambda}} \geq \max_{x\in\{a,b\}}\left[\frac{1}{n}\sum_{i=1}^{n}|x-x_i|^{\lambda}\right]^{\frac{1}{\lambda}}$ $> \frac{1}{2}(b-a)$ .

## Theorem 3.1.5

For $\lambda = 1$ , Chebyshev points of order $n$ , $n > 1$ for $E$ are not unique. Indeed, any set of points of the form

$$\left\{\underbrace{\frac{b+a}{2}+c,\frac{b+a}{2}+c,\dots,\frac{b+a}{2}+c}_{\frac{n}{2}}\;;\;\underbrace{\frac{b+a}{2}-c,\dots,\frac{b+a}{2}-c}_{\frac{n}{2}}\right\} \quad \text{for } n \text{ even}$$

and $|c| \leq \frac{1}{2}(b-a)$ is a set of Chebyshev points of order $n$ for $E$.

If $n$ is odd, any set of the form

$$\left\{\underbrace{\frac{b+a}{2}+c,\dots,\frac{b+a}{2}+c}_{\frac{n-1}{2}}\;;\;\frac{b+a}{2};\underbrace{\frac{b+a}{2}-c,\dots,\frac{b+a}{2}-c}_{\frac{n-1}{2}}\right\} \quad \text{where again } |c| \leq \tfrac{1}{2}(b-a)$$

is a set of Chebyshev points of order $n$ for $E$.

<u>Proof</u>: For $n$ even:

$$\max \frac{1}{n}\sum_{i=1}^{n} |x-x_i| = \frac{1}{n}\left[\frac{n}{2}\left[b-\left(\frac{b+a}{2}+c\right)\right]+\frac{n}{2}\left[b-\left(\frac{b+a}{2}-c\right)\right]\right] = \tfrac{1}{2}(b-a) = \tau_n^{(1)}(E) .$$

For $n$ odd:

$$\max \frac{1}{n}\sum_{i=1}^{n} |x-x_i| = \frac{1}{n}\left[\frac{n-1}{2}\left[b-\left(\frac{b+a}{2}+c\right)\right]+\frac{n-1}{2}\left[b-\left(\frac{b+a}{2}-c\right)\right]+b-\left(\frac{b+a}{2}\right)\right]$$

$$= \tfrac{1}{2}(b-a) = \tau_n^{(1)}(E) .$$

For $\lambda > 1$, the Chebyshev function of order $n$ for $E$ is unique. It is

$$C_n^{(\lambda)}(x;E) = n^{\frac{1}{\lambda}} \left|x-\left(\frac{b+a}{2}\right)\right| \tag{3.1.22}$$

From Theorem 2.3.1 and Definition 2.3.4 we see that

$$m_n^{(\lambda)}(E) = n^{\frac{1}{\lambda}} \tau_n^{(\lambda)}(E) = n^{\frac{1}{\lambda}} \tfrac{1}{2}(b-a) \quad \text{and}$$

$$\mathcal{L}_{C_n^{(\lambda)}}\{x \in R_1 \mid |x-\left(\frac{b+a}{2}\right)| \leq \tfrac{1}{2}(b-a)\} = [a,b] \tag{3.1.23}$$

$$\hat{\mathcal{L}}_{C_n^{(\lambda)}} : \{x \in R_1 \mid |x - \left(\tfrac{b+a}{2}\right)| = \tfrac{1}{2}(b-a)\} = \{a,b\} \qquad (3.1.24)$$

For $\lambda = 1$ , $n$ even, Chebyshev functions of order $n$ for $E$ take the form

$$C_n^{(1)}(x;E) = \frac{n}{2}\left[\left|x - \left(\frac{b+a}{2} + c\right)\right| + \left|x - \left(\frac{b+a}{2} - c\right)\right|\right] \qquad (3.1.25)$$

where $|c| \le \frac{1}{2}(b-a)$ . Thus

$$\mathcal{L}_{C_n^{(1)}} : \{x \in R_1 \mid \left|x - \left(\frac{b+a}{2} + c\right)\right| + \left|x - \left(\frac{b+a}{2} - c\right)\right| \le (b-a)\} = [a,b] \qquad (3.1.26)$$

$$\hat{\mathcal{L}}_{C_n^{(1)}} : \{x \in R_1 \mid \left|x - \left(\frac{b+a}{2} + c\right)\right| + \left|x - \left(\frac{b+a}{2} - c\right)\right| = (b-a)\}$$

$$= \begin{Bmatrix} \{a,b\} & \text{for} & c < \frac{1}{2}(b-a) \\ [a,b] & \text{for} & c = \frac{1}{2}(b-a) \end{Bmatrix} \qquad (3.1.27)$$

For $\lambda = 1$ , $n$ odd, Chebyshev functions of order $n$ for $E$ take the form

$$C_n^{(1)}(x;E) = \frac{n-1}{2}\left[\left|x - \left(\frac{b+a}{2} + c\right)\right| + \left|x - \left(\frac{b+a}{2} - c\right)\right|\right] + \left|x - \frac{b+a}{2}\right| \qquad (3.1.28)$$

where $|c| \le \frac{1}{2}(b-a)$ . Thus

$$\mathcal{L}_{C_n^{(1)}} : \{x \in R_1 \mid \left(\frac{n-1}{2}\right)\left[\left|x - \left(\frac{b+a}{2} + c\right)\right| + \left|x - \left(\frac{b+a}{2} - c\right)\right|\right] + \left|x - \left(\frac{b+a}{2}\right)\right| \qquad (3.1.29)$$

$$\le \frac{n}{2}(b-a)\} = [a,b] .$$

$$\hat{\mathcal{L}}_{C_n^{(1)}} : \{x \in R_1 \,|\, \binom{n-1}{2} \left[ \left|x - \left(\frac{b+a}{2} + c\right)\right| + \left|x - \left(\frac{b+a}{2} - c\right)\right| \right] + \left|x - \left(\frac{b+a}{2}\right)\right|$$

$$= \frac{n}{2} (b - a)\} = \{a,b\} . \qquad (3.1.30)$$

Theorem 3.1.6

For $\lambda = 1$ , $\chi^{(1)}(E) = \delta^{(1)}(E)$ and for $\lambda > 1$ , $\chi^{(\lambda)}(E) < \delta^{(\lambda)}(E)$ for any compact set $E$ in $R_1$ .

Proof: From Corollary 3.1.1 and Theorem 3.1.3 we have

$\delta^{(\lambda)}(E) = \frac{1}{2^{\frac{1}{\lambda}}} (b - a) \geq \frac{1}{2} (b - a) = \chi^{(\lambda)}(E)$ with equality if and only if $\lambda = 1$ .

The (generalized) restricted Chebyshev problem for a compact set $E$ in $R_1$ is more difficult and only partial results have been obtained.

Theorem 3.1.7

For $\lambda = 1$ , $\tilde{\chi}^{(1)}(E) = \chi^{(1)}(E) = \frac{1}{2}(b - a)$ .

Proof: Since there exists a set of Chebyshev points of order $n$ , $n$ even, for $E$ which lie in $E$ , namely $\{\underbrace{a,a,\ldots,a}_{\frac{n}{2}} ; \underbrace{b,b,\ldots,b}_{\frac{n}{2}}\}$ by Theorem 2.5.2 this set is a set of restricted Chebyshev points of order $n$ for $E$. Thus for $n$ even, $\tilde{\tau}_n^{(1)}(E) = \max_{x \in E} \frac{1}{n}\left[ \frac{n}{2} |x-a| + \frac{n}{2} |x-b| \right] = \frac{1}{2}(b - a)$ and $\tilde{\chi}^{(1)}(E) = \frac{1}{2}(b - a)$ .

We note that for $\lambda = 1$ the restricted Chebyshev points for $E$

may or may not be unique. The points are unique if and only if $E$ does not contain the point $\frac{a+b}{2}$ nor any pair of points symmetrical with respect to $\frac{a+b}{2}$ except for the points $a$ and $b$.

Example 3.1.1

Let $E = \{a,b\}$. Then for $\lambda = 1$ the generalized restricted Chebyshev points of order $n$, $n$ even, for $E$ are unique, but the generalized Chebyshev points of order $n$, $n$ even, for $E$ are not unique. The restricted Chebyshev points are $\{\underbrace{a,a,\ldots,a}_{\frac{n}{2}}\ ;\ \underbrace{b,b,\ldots,b}_{\frac{n}{2}}\}$ ;

the Chebyshev points are $\left\{\underbrace{\frac{b+a}{2} + c, \frac{b+a}{2} + c, \ldots, \frac{b+a}{2} + c}_{\frac{n}{2}} ; \underbrace{\frac{b+a}{2} - c, \ldots, \frac{b+a}{2} - c}_{\frac{n}{2}}\right\}$

where $|c| \le \frac{1}{2}(b-a)$.

For $\lambda > 1$, if $\frac{a+b}{2} \in E$, then we have $\tilde{\chi}^{(\lambda)}(E) = \frac{1}{2}(b-a)$ from Theorem 2.5.2. If $\frac{a+b}{2} \notin E$, then we do know for $n$ even

$$\tilde{T}_n^{(\lambda)}(E) \le \max_{x\in E} \left(\frac{1}{n}\left[\frac{n}{2}\,|x-a|^{\lambda} + \frac{n}{2}\,|x-b|^{\lambda}\right]\right)^{\frac{1}{\lambda}} \le \frac{1}{2^{\frac{1}{\lambda}}}(b-a) \text{ . Thus}$$

$\tilde{\chi}^{(\lambda)}(E) \le \frac{1}{2^{\frac{1}{\lambda}}}(b-a)$ and since $\chi^{(\lambda)}(E) \le \tilde{\chi}^{(\lambda)}(E)$ we have

$$\frac{1}{2}(b-a) \le \tilde{\chi}^{(\lambda)}(E) \le \frac{1}{2^{\frac{1}{\lambda}}}(b-a) \qquad (3.1.31)$$

Example 3.1.1 above provides us with a case where $\chi(E) < \tilde{\chi}(E)$. For if $\lambda > 1$, then $\chi^{(\lambda)}(E) = \frac{1}{2}(b-a)$ and $\tilde{\chi}^{(\lambda)}(E) = \frac{1}{2^{\frac{1}{\lambda}}}(b-a)$.

One can see that the restricted problem is very "delicate," for if we introduce just the single point $\frac{a+b}{2}$ into the set E of Example 3.1.1, then $\tilde{\chi}^{(\lambda)}(E)$ for $\lambda > 1$ goes from its upper bound to its lower bound.

## 3.2 $R_2$ Two-dimensional Real Euclidean Space

Although we will write $R_2$, the results for this section are valid for the complex plane as well. Of course for $\lambda = 0$, the study of the topics we will discuss comprises the classical theory outlined in Chapter 1. The case of $\lambda \geq 1$ in $R_2$ (or the complex plane) is also of interest since we shall show that one may calculate $\chi^{(\lambda)}(E)$ and geometrically characterize the Chebyshev points of E for all compact sets in $R_2$. This will be shown in the main theorem of section 3.2. To develop the techniques for the proof of this theorem we are led to a consideration of two topics, seemingly unrelated to the present field of inquiry. One concerns some geometric properties of point sets and the other deals with a problem known in the literature as Steiner's problem. (We ask the reader to be indulgent for these necessary sidetrips.)

The importance of the role of points a and b in the case of a set E in $R_1$ suggests that one consider some "bounding" type of set for E in $R_2$.

Definition 3.2.1

Let E be a compact [1] set in $R_2$. The circle of smallest radius which contains E will be called the spanning circle of E and will

---

[1] This definition as well as Theorems 3.2.1 and 3.2.2 are valid for point sets which are merely bounded, but we shall always restrict our discussion to compact sets.

be denoted by $C_E$ .

Theorem 3.2.1

The spanning circle of $E$ is unique, and it either 1) contains two boundary points of $E$ which are at the ends of a diameter of the circle, or 2) contains three boundary points of $E$ which form an acute triangle.

Theorem 3.2.1 may be found as a problem (with solution) in Yaglom [33]. It enables us to decompose the set of all compact sets in $R_2$ in the following manner:

Definition 3.2.2

Compact sets $E$ in $R_2$ such that $E \cap C_E$ contains points $P, P'$, where the segment $PP'$ is a diameter of $C_E$ shall be called Case I sets. All other compact sets $E$ of $R_2$ shall be called Case II sets.

We note that from Theorem 3.2.1, the intersection of a Case II set with its spanning circle contains the vertices of an acute triangle.

Since the set $E$ is enclosed by $C_E$ , we must have $d(E) \leq d(C_E)$ where as before $d(E)$ denotes the topological diameter of $E$ . If $R$ is the radius of $C_E$ , then we have

$$d(E) \leq 2R \qquad (3.2.1)$$

with equality if and only if $E$ is a Case I set. Equation (3.2.1) provides us with a lower bound on the radius of the spanning circle. For an upper bound we have the theorem of H.W.E. Jung [17]:

Theorem 3.2.2

Let $E$ be a compact set in $R_2$ of topological diameter $d(E)$ .

Then $C_E$ has a radius $R$ where $R \leq \frac{d(E)\sqrt{3}}{3}$ .

<u>Proof</u>: For Case I sets: $d(E) = 2R$ . Thus $R = \frac{d(E)}{2} < \frac{d(E)\sqrt{3}}{3}$

For Case II sets: At least one angle of the acute triangle whose vertices lie in $E \cap C_E$ must be greater than or equal to $\frac{\pi}{3}$ . Let us denote such an angle by $\sphericalangle A$. Then $\sin A \geq \frac{\sqrt{3}}{2}$ . Now $2R = \frac{a}{\sin A}$ and $a < d(E)$ . Thus $R \leq \frac{d(E) \cdot 2}{2\sqrt{3}} = \frac{d(E)\sqrt{3}}{3}$ .

We note in passing that this bound cannot be improved, for if $E$ is an equilateral triangle then its spanning circle coincides with its circumscribing circle of radius $R = \frac{d(E)\sqrt{3}}{3}$ . The coincidence of the spanning and circumscribing circles holds for all acute and right triangles. However for obtuse triangles, the spanning and circumscribing circles do not coincide. The obtuse triangle is an example of a Case I set where the diameter of the spanning circle is equal to the longest side of the triangle.

Theorem 3.2.2 and equation (3.2.1) give

$$\frac{d(E)}{2} \leq R \leq \frac{d(E)\sqrt{3}}{3} \tag{3.2.2}$$

Although its relevance is not discernible at this time, we shall now discuss a problem known as Steiner's problem:

Let $A_1$, $A_2$, and $A_3$ be 3 points in the plane. Find a point $Q$ in the plane such that $\sum_{i=1}^{3} |A_i - Q|$ shall be minimal, where $|A_i - Q|$ denotes the distance from the point $A_i$ to the point $Q$ .

The solution of this problem is discussed in Courant and Robbins [3]. It is shown there that if the triangle formed by $A_1$, $A_2$, and $A_3$

contains no vertex angle greater than 120 degrees, then $Q$ is the point within the triangle such that the lines $A_1Q$ , $A_2Q$ , and $A_3Q$ form angles of 120 degrees with each other. If the triangle formed by $A_1$, $A_2$, and $A_3$ contains a vertex angle greater than or equal to 120 degrees, then $Q$ coincides with this vertex. The point $Q$ is called the Steiner point of the triangle.

An extended Steiner problem is discussed in Steinhaus [29]. Let $w_{A_i}$ be a non-negative weight attached to $A_i$ , $i = 1,2,3$ . Find a point $Q$ in the plane such that $\sum_{i=1}^{3} w_{A_i} |A_i - Q|$ shall be minimal. One can associate with this problem the following statics problem. Let us "place" triangle $A_1A_2A_3$ on a table and drill holes where the vertices are located. We attach strings to the weights $w_{A_1}$, $w_{A_2}$, and $w_{A_3}$ below the table, pass the strings through the holes at $A_1$, $A_2$, and $A_3$ and join them in a knot. The point at which the knot comes to rest (at which the system is in equilibrium) is the desired point $Q$ . At $Q$ the sum of the forces equals zero. This sum can be represented by a closed polygon - in this case a triangle with sides $w_{A_1}$, $w_{A_2}$, and $w_{A_3}$ . Let $\alpha_3$ be the exterior angle between the sides $w_{A_1}$ and $w_{A_2}$ ; $\alpha_1$ be the exterior angle between the sides $w_{A_2}$ and $w_{A_3}$ ; and $\alpha_2$ be the exterior angle between the sides $w_{A_3}$ and $w_{A_1}$ . Then the Steiner point $Q$ is the point such that $\alpha_3 = \sphericalangle A_1QA_2$ , $\alpha_1 = \sphericalangle A_2QA_3$, and $\alpha_2 = \sphericalangle A_3QA_1$ . (If the weights are such that one weight, say $w_{A_1}$ is so great that the other two are not able to withstand its pull, then the knot will be caught by the hole at $A_1$ and $Q = A_1$ .)

Our concern will be a problem which, in part, could be considered a reverse of the extended problem above. Only the case where $A_1$, $A_2$, and $A_3$ form an acute triangle will ultimately be of interest to us, so we shall restrict our consideration to the question:

If $A_1$, $A_2$, and $A_3$ are the vertices of an acute triangle, can we attach non-negative weights $w_{A_1}$, $w_{A_2}$, and $w_{A_3}$ to the points $A_1$, $A_2$, and $A_3$ to make any given point $Q$ in the interior of triangle $A_1A_2A_3$ a " $\lambda$-Steiner point;" that is, to minimize the function

$$\sum_{i=1}^{3} w_{A_i} |A_i - Q|^{\lambda} \tag{3.2.3}$$

for $\lambda \geq 1$ ? (For $\lambda = 1$ , $Q$ would be a Steiner point in the sense of the extended problem discussed in Steinhaus [29].)

For $\lambda = 1$ , the case previously described provides the solution. If

$$\alpha_3 = \sphericalangle A_1QA_2 \ , \ \alpha_1 = \sphericalangle A_2QA_3 \ , \text{ and } \ \alpha_2 = \sphericalangle A_3QA_1 \tag{3.2.4}$$

we draw a triangle with exterior angles $\alpha_3$, $\alpha_1$, and $\alpha_2$ . The sides of the triangle will correspond to the desired weights. Since $\sin \theta = \sin(\pi - \theta)$ , we have

$$\frac{w_{A_1}}{\sin \alpha_1} = \frac{w_{A_2}}{\sin \alpha_2} = \frac{w_{A_3}}{\sin \alpha_3} \tag{3.2.5}$$

and since $Q$ is an interior point of the triangle, all the weights will be strictly positive.

We note that the solution, of course, is not unique, but since all triangles with given angles are similar, all sets of solutions will be proportional.

For $\lambda > 1$, we will transpose our problem into one in approximation theory. Let $E$ be a compact set in the complex plane and let $\mu$ be a positive measure on $\mathbb{C}$. We denote the identity function by $I(z)$ and ask, "What is the best approximation in p-norm, $p > 1$ to $I(z)$ on $E$ from the subspace consisting of the constant functions?" Thus, we are asked to find some constant $Q$ such that

$$\int |I(z) - X|^p \, d\mu \geq \int |I(z) - Q|^p \, d\mu \qquad (3.2.6)$$

for all constants $X$.

Now, if we choose $E$ to be the finite set of points $\{A_1, A_2, A_3\}$ and choose the measure $\mu$ to assign the weights $w_{A_i}$ at $A_i$, $i = 1,2,3$ then we see that setting $p = \lambda$, the right hand side of (3.2.6) is identical to the function (3.2.3) under consideration in the reverse $\lambda$-Steiner problem.

<u>Definition 3.2.3</u>

If $z \in \mathbb{C}$, we define $\operatorname{sgn} z = \begin{cases} 0 & \text{if } z = 0 \\ \frac{z}{|z|} & \text{if } z \neq 0 \end{cases}$ and $\overline{\operatorname{sgn}}\, z = \operatorname{sgn} \bar{z}$.

Looking at the Steiner problem in terms of an approximation problem enables one to use the following theorem.

<u>Theorem 3.2.3</u> [1]

A sufficiency condition for the best approximation in the case $p > 1$ is given by

---

[1] The condition stated is both necessary and sufficient for the best approximation, but we use only the sufficiency.

$$\sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p-1}\ \overline{\mathrm{sgn}}\ w_{A_i} (A_i - Q) X = 0 \quad \text{for every} \quad X \quad \text{in} \quad ¢ \ .$$

<u>Proof:</u>

$$\sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p} = \sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p-1} (A_i - Q)\ \overline{\mathrm{sgn}}\ (A_i - Q)$$

$$= \sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p-1} (A_i - X)\ \overline{\mathrm{sgn}}\ (A_i - Q)$$

$$+ \sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p-1} (X - Q)\ \overline{\mathrm{sgn}}\ (A_i - Q) \ .$$

The second term is zero by hypothesis since $\overline{\mathrm{sgn}}\ (A_i - Q) = \overline{\mathrm{sgn}}\ w_{A_i} (A_i - Q)$ and $X - Q \in$ ¢ . Thus

$$\sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p} \le \sum_{i=1}^{3} w_{A_i}^{\frac{1}{q}} |A_i - Q|^{p-1} \left(w_{A_i}^{\frac{1}{p}} |A_i - X|\right) \text{ where } \frac{1}{q} = 1 - \frac{1}{p} \ .$$

Applying Holder's inequality this gives

$$\sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p} \le \left(\sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p}\right)^{\frac{1}{q}} \left(\sum_{i=1}^{3} w_{A_i} |A_i - X|^{p}\right)^{\frac{1}{p}} \quad \text{where}$$

$\frac{1}{p} + \frac{1}{q} = 1$ . Thus

$$\left(\sum_{i=1}^{3} w_{A_i} |A_i - Q|^{p}\right)^{\frac{1}{p}} \le \left(\sum_{i=1}^{3} w_{A_i} |A_i - X|^{p}\right)^{\frac{1}{p}}$$

and raising both sides to the p'th power gives the desired result.

<u>Theorem 3.2.4</u>

$w_{A_i} > 0$ , $i = 1,2,3$ may be chosen so as to satisfy the sufficiency condition in Theorem 3.2.3.

Proof: Since $Q$ is in triangle $A_1A_2A_3$ we may write $Q$ using its barycentric coordinates relative to the points $A_1, A_2$, and $A_3$. Thus $Q = \sum_{i=1}^{3} \lambda_i A_i$ with $\sum_{i=1}^{3} \lambda_i = 1$ and since $Q$ is in the interior of $\Delta A_1A_2A_3$, $\lambda_i > 0$, $i = 1,2,3$. Hence $\sum_{i=1}^{3} \lambda_i (A_i - Q) = 0$ and also $\sum_{i=1}^{3} \lambda_i \overline{(A_i - Q)} = 0$. Set

$$w_{A_i} = \lambda_i |A_i - Q|^{2-p} \quad \text{for} \quad i = 1,2,3. \tag{3.2.7}$$

Then

$$0 = \sum_{i=1}^{3} \lambda_i \overline{(A_i-Q)} = \sum_{i=1}^{3} w_{A_i} |A_i-Q|^{p-2} \overline{(A_i-Q)} = \sum_{i=1}^{3} w_{A_i} |A_i-Q|^{p-1} \overline{\operatorname{sgn}}(A_i-Q)$$

Hence

$$\sum_{i=1}^{3} w_{A_i} |A_i-Q|^{p-1} \overline{\operatorname{sgn}}(A_i-Q)\, X = 0 \quad \text{for all} \quad X \in \mathbb{C}.$$

Thus we have found a set of positive weights given by (3.2.7) that minimize the function (3.2.3).

We now have the machinery to prove the main theorem of this section.

To simplify the notation we will designate points in $R_2$ by using complex numbers. Thus, if $P$ and $Q$ are points in $R_2$, we shall write $|P-Q|$ for the distance between $P$ and $Q$.

Theorem 3.2.5

Let $E$ be any compact set in $R_2$. Let $C_E$ designate the spanning circle of $E$. We shall denote the center of $C_E$ by $O'$ and its radius by $R$. Then $\chi^{(\lambda)}(E) = R$ for $\lambda \geq 1$.

Proof: Let $E$ be a Case I set. Then $E \cap C_E \supset \{P, P'\}$ where $|P-P'| = 2R$

and the point $O'$ is the mid-point of segment $PP'$ .

Let $z_i = O'$ for $i = 1,2,\ldots,n$ . Then

$$T_n^{(\lambda)}(E) \le \max_{z\in E}\left(\frac{1}{n}\sum_{i=1}^{n}|z-z_i|^{\lambda}\right)^{\frac{1}{\lambda}} = \left[\frac{1}{n}\,(n)\,|P-O'|^{\lambda}\right]^{\frac{1}{\lambda}} = R\ .$$

Now let $z_i$ , $i = 1,2,\ldots,n$ be arbitrary points of $R_2$ . For $\lambda \ge 1$ , Minkowski's inequality gives

$$n^{\frac{1}{\lambda}}\,|P-P'| = \left(\sum_{i=1}^{n}|P-z_i + z_i-P'|^{\lambda}\right)^{\frac{1}{\lambda}} \le \left(\sum_{i=1}^{n}\left(|P-z_i| + |z_i-P'|\right)^{\lambda}\right)^{\frac{1}{\lambda}}$$

$$\le \left(\sum_{i=1}^{n}|P-z_i|^{\lambda}\right)^{\frac{1}{\lambda}} + \left(\sum_{i=1}^{n}|z_i-P'|^{\lambda}\right)^{\frac{1}{\lambda}}\ .$$

Hence either $\left(\sum_{i=1}^{n}|P-z_i|^{\lambda}\right)^{\frac{1}{\lambda}} \ge \frac{1}{2}\,n^{\frac{1}{\lambda}}|P-P'|$ or $\left(\sum_{i=1}^{n}|P'-z_i|^{\lambda}\right)^{\frac{1}{\lambda}} \ge \frac{1}{2}\,n^{\frac{1}{\lambda}}|P-P'|$ .

Thus if $z_i^*$ , $i = 1,2,\ldots,n$ is a set of Chebyshev points of order $n$ for $E$ ,

$$T_n^{(\lambda)}(E) = \max_{z\in E}\left(\frac{1}{n}\sum_{i=1}^{n}|z-z_i^*|^{\lambda}\right)^{\frac{1}{\lambda}} \ge \max_{z\in\{P,P'\}}\left(\frac{1}{n}\sum_{i=1}^{n}|z-z_i^*|^{\lambda}\right)^{\frac{1}{\lambda}}$$

$$\ge \left(\frac{1}{n^{\frac{1}{\lambda}}}\,\frac{1}{2}\,n^{\frac{1}{\lambda}}|P-P'|\right) = \frac{1}{2}(2R) = R\ .$$

Hence $T_n^{(\lambda)}(E) = R$ for every $n$ and $\chi^{(\lambda)}(E) = R$ .

Now let $E$ be a Case II set. Then $E \cap C_E \supset \{A_1,A_2,A_3\}$ where $\Delta\, A_1A_2A_3$ is acute. Again placing $z_i = O'$ for $i = 1,2,\ldots,n$ gives $T_n^{(\lambda)}(E) \le R$ . Our excursion into the Steiner problem helps us to establish the reverse inequality.

Suppose $z_i^*$ , $i = 1,2,\ldots,n$ is a set of Chebyshev points of order $n$ for $E$ . Then

$$T_n^{(\lambda)}(E) = \max_{z\in E}\left(\frac{1}{n}\sum_{i=1}^{n} |z - z_i^*|^{\lambda}\right)^{\frac{1}{\lambda}} \geq \max_{z\in\{A_1,A_2,A_3\}}\left(\frac{1}{n}\sum_{i=1}^{n} |z - z_i^*|^{\lambda}\right)^{\frac{1}{\lambda}} \tag{3.2.8}$$

The maximum of a set of numbers is greater than or equal to any average function of them by (2.1.3). We shall use a weighted $\lambda$-th power average as in Definition 2.1.5. Thus

$$\max_{z\in\{A_1,A_2,A_3\}}\left(\frac{1}{n}\sum_{i=1}^{n} |z - z_i^*|^{\lambda}\right)^{\frac{1}{\lambda}} \geq$$

$$\left[\frac{w_{A_1}\left[\left(\frac{1}{n}\sum_{i=1}^{n}|A_1-z_i^*|^{\lambda}\right)^{\frac{1}{\lambda}}\right]^{\lambda} + w_{A_2}\left[\left(\frac{1}{n}\sum_{i=1}^{n}|A_2-z_i^*|^{\lambda}\right)^{\frac{1}{\lambda}}\right]^{\lambda} + w_{A_3}\left[\left(\frac{1}{n}\sum_{i=1}^{n}|A_3-z_i^*|^{\lambda}\right)^{\frac{1}{\lambda}}\right]^{\lambda}}{w_{A_1} + w_{A_2} + w_{A_3}}\right]^{\frac{1}{\lambda}} \tag{3.2.9}$$

Rearranging terms, the right hand side of (3.2.9) becomes

$$\left[\frac{\frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-z_1^*|^{\lambda} + \frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-z_2^*|^{\lambda} + \ldots + \frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-z_n^*|^{\lambda}}{w_{A_1} + w_{A_2} + w_{A_3}}\right]^{\frac{1}{\lambda}} \tag{3.2.10}$$

Since $\Delta\, A_1A_2A_3$ is acute, the point $O'$ is properly contained within the triangle and its barycentric coordinates $\lambda_j$ are strictly greater than zero for $j = 1,2,3$. Our weighted average is chosen with weights $w_{A_j}$, $j = 1,2,3$ such that the point $O'$ is the " $\lambda$-Steiner point", $\lambda \geq 1$, in accordance with (3.2.7) with $Q = O'$. Thus (3.2.10) is greater than or equal to

$$\left[\frac{\frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-O'|^{\lambda} + \frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-O'|^{\lambda} + \dots + \frac{1}{n}\sum_{j=1}^{n} w_{A_j}|A_j-O'|^{\lambda}}{\sum_{j=1}^{3} w_{A_j}}\right]^{\frac{1}{\lambda}}$$

$$= \left[\frac{\frac{1}{n}\, n \sum_{j=1}^{3} w_{A_j}|A_j-O'|^{\lambda}}{\sum_{j=1}^{3} w_{A_j}}\right]^{\frac{1}{\lambda}} = \left[\frac{\frac{1}{n}(n)\, R^{\lambda} \sum_{j=1}^{3} w_{A_j}}{\sum_{j=1}^{3} w_{A_j}}\right]^{\frac{1}{\lambda}} = R\,.$$

Hence $T_n^{(\lambda)}(E) = R$ for every $n$ and $\chi^{(\lambda)}(E) = R$ for all Case II sets $E$ in $R_2$. Thus the result is established for all compact sets $E$ in $R_2$.

We note that the procedure describing how to choose weights so as to force any particular point $Q$ to be the " $\lambda$-Steiner point" can be extended to the case where $A_1, A_2, \dots, A_n$ are the vertices of a closed convex $n$-gon and $Q$ is some interior point of the $n$-gon. However, such a convex body is a compact set and Theorem 3.2.5 already gives us the information we want regarding its Chebyshev constant.

Theorem 3.2.6

For $\lambda > 1$ , the Chebyshev points of order $n$ for $E$ are unique and are $z_i^* = 0'$ for $i = 1,2,\dots,n$ .

Proof: We note from Theorem 3.2.5 that the set $z_i^* = 0'$ for $i = 1,2,\dots,n$ is a set of Chebyshev points of order $n$ for $E$ . The approximation problem, being a problem in a finite dimensional space with p-norm, $p = \lambda > 1$ admits a unique solution. Thus there exist weights $w_{A_1}$, $w_{A_2}$, $w_{A_3}$ , such that $\sum_{j=1}^{3} w_{A_j} |A_j - X|^{\lambda} > \sum_{j=1}^{3} w_{A_j} |A_j - 0'|^{\lambda}$ for $X \neq 0'$ .

Hence for arbitrary $z_i$ , $i = 1,2,\dots,n$ $z_i \neq 0'$ for every $i$ , we have

$$\max_{z\in E}\left[\frac{1}{n}\sum_{i=1}^{n}|z-z_i|^{\lambda}\right]^{\frac{1}{\lambda}} \ge \max_{z\in\{A_1,A_2,A_3\}}\left[\frac{1}{n}\sum_{i=1}^{n}|z-z_i|^{\lambda}\right]^{\frac{1}{\lambda}}$$

$$\ge \left[\frac{w_{A_1}\left(\frac{1}{n}\sum_{i=1}^{n}|A_1-z_i|^{\lambda}\right) + w_{A_2}\left(\frac{1}{n}\sum_{i=1}^{n}|A_2-z_i|^{\lambda}\right) + w_{A_3}\left(\frac{1}{n}\sum_{i=1}^{n}|A_3-z_i|^{\lambda}\right)}{\sum_{j=1}^{3} w_{A_j}}\right]^{\frac{1}{\lambda}}$$

$$= \left[\frac{\frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-z_1|^{\lambda} + \frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-z_2|^{\lambda} + \dots + \frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-z_n|^{\lambda}}{\sum_{j=1}^{3} w_{A_j}}\right]^{\frac{1}{\lambda}}$$

$$> \left[\frac{\frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-0'|^{\lambda} + \frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-0'|^{\lambda} + \dots + \frac{1}{n}\sum_{j=1}^{3} w_{A_j}|A_j-0'|^{\lambda}}{\sum_{j=1}^{3} w_{A_j}} \qquad (3.2.11)\right]^{\frac{1}{\lambda}}$$

$= R$ .

For $\lambda = 1$, the best approximation problem does not necessarily yield a unique solution, so we must examine that case separately.

Theorem 3.2.7

For $\lambda = 1$, the Chebyshev points of order $n$ for $E$ are unique if $|E \cap C_E| \geq 3$.

Proof: We note first that the condition is satisfied for all Case II sets since $E \cap C_E \supset \{A_1, A_2, A_3\}$ where $\Delta\, A_1A_2A_3$ is acute. It is also satisfied for those Case I sets such that $E \cap C_E \supset \{A_1, A_2, A_3\}$ where $\Delta\, A_1A_2A_3$ is a right angle.

For the case $\lambda = 1$, we found weights minimizing (3.2.3) using a geometric argument. One could also solve this problem using a sufficiency condition for best approximation in the $p = 1$ case. The sufficiency condition for $\sum_{i=1}^{3} w_{A_i} |A_i - Q| \leq \sum_{i=1}^{3} w_{A_i} |A_i - X|$ is

$$\sum_{i=1}^{3} w_{A_i} X\, \overline{\mathrm{sgn}}(A_i - Q) = 0 \tag{3.2.12}$$

for all $X$ in $\mathbb{C}$, and it is proved along the lines of Theorem 3.2.3. Since

$$\sum_{i=1}^{3} w_{A_i} |A_i - Q| = \sum_{i=1}^{3} w_{A_i} (A_i - X)\, \overline{\mathrm{sgn}}\, (A_i - Q) \tag{3.2.13}$$

we must have

$$\sum_{i=1}^{3} \arg[(A_i - X)\, \overline{\mathrm{sgn}}(A_i - Q)] = 0 \tag{3.2.14}$$

Now

$$\sum_{i=1}^{3} w_{A_i}(A_i-X)\,\overline{\operatorname{sgn}}(A_i-Q) \le \left|\sum_{i=1}^{3} w_{A_i}(A_i-X)\,\overline{\operatorname{sgn}}(A_i-Q)\right| \le \sum_{i=1}^{3} w_{A_i}|A_i-X| .$$

The last inequality on the right is strict unless either 1) $A_i-X = 0$ for all $i$ or 2) $\arg[(A_i-X)\,\overline{\operatorname{sgn}}(A_i-Q)]$ is the same for all $i$. Possibility 1) can never happen since this would imply that $A = (A_1,A_2,A_3)$ is the same vector as $\hat{X} = (X,X,X)$ which cannot be, since the $A_i$'s are distinct points. In view of (3.2.14) possibility 2) states that $\arg[(A_i-X)\,\overline{\operatorname{sgn}}(A_i-Q)] = 0$ for all $i$. Thus,

$$\arg(A_i - X) + \arg\overline{\operatorname{sgn}}(A_i - Q) = 0 \qquad \text{or}$$

$$\arg(A_i - X) - \arg(A_i - Q) = 0 . \quad \text{Hence,} \quad \arg(A_i - X) = \arg(A_i - Q) .$$

Thus there exists $\mu \neq 0$ such that $(A_i - X) = \mu(A_i - Q)$. Substituting in (3.2.13) we have $\sum_{i=1}^{3} w_{A_i}|A_i - Q| = \sum w_{A_i}\mu(A_i - Q)\,\overline{\operatorname{sgn}}(A_i - Q)$

or $\sum_{i=1}^{3} w_{A_i}|A_i - Q| = \mu \sum_{i=1}^{3} w_{A_i}|A_i - Q|$. Hence $\mu = 1$ and $A_i - Q = A_i - X$ or $X = Q$. Using the fact that $O'$ is in the convex hull of $\Delta A_1A_2A_3$ we can attach weights $w_{A_1}$, $w_{A_2}$, and $w_{A_3}$ to $A_1, A_2$, and $A_3$ such that

$$\sum_{i=1}^{3} w_{A_i}|A_i - X| > \sum_{i=1}^{3} w_{A_i}|A_i - O'| . \qquad (3.2.15)$$

(If $\Delta A_1A_2A_3$ is a right triangle we attach the weight zero to the vertex of the right angle.) If one repeats the proof of Theorem (3.2.6) with $\lambda = 1$ condition (3.2.15) shows that the inequality in line (3.2.11) remains strict. Thus the Chebyshev points of order $n$ for those sets described in the hypothesis are $z_i = O'$ for $i = 1,2,\ldots,n$ and this set is unique.

We now turn our attention to the case where $|E \cap C_E| = 2$. Thus $E \cap C_E = \{P,P'\}$.

<u>Definition 3.2.4</u>

Let $\mathcal{E}$ be the set of ellipses with major axis $PP'$. Let $M_E \in \mathcal{E}$ be the ellipse of largest eccentricity which encloses $E$. We shall call $M_E$ the <u>spanning ellipse of $E$</u>.

If we denote the eccentricity of $M_E$ by $\epsilon$, we note that if $|E \cap C_E| = 2$, then $\epsilon > 0$. For those sets of Case I such that $|E \cap C_E| \geq 3$ we have $\epsilon = 0$ and $M_E = C_E$.

<u>Theorem 3.2.8</u>

For $\lambda = 1$, the Chebyshev points of order $n$, $n > 1$ for $E$ are not unique if $|E \cap C_E| = 2$.

<u>Proof</u>: Let $M_E$ be the spanning ellipse of $E$ with eccentricity $\epsilon$. Let $F$ and $F'$ be the foci of $M_E$. The length of the segments $O'F$ and $O'F'$ is $c = R\epsilon$. If $n$ is even we let

$$z_i = \begin{cases} O' + d & i = 1,2,\ldots,\frac{n}{2} \\ O' - d & i = \frac{n}{2}+1,\ldots,n \end{cases} \tag{3.2.16}$$

where the point $O' + d$ is a translation of the point $O'$ <u>along the major axis $PP'$</u>. Then

$$\frac{1}{n}\max_{z\in E}\sum_{i=1}^{n}|z-z_i| = \frac{1}{n}\max_{z\in E}\left[\frac{n}{2}|z-(O'+d)| + \frac{n}{2}|z-(O'-d)|\right].$$

Now the ellipse with foci at $O'+d$ and $O'-d$ for $d \leq c$ has eccentricity smaller than or at most equal (in the case $d = c$) to the eccentricity of $M_E$. Thus if $d < c$, then for every $z$ in $E$ $z \neq P,P'$ we have

$$|z-(O'+d)| + |z-(O'-d)| < 2R \quad \text{and}$$
$$|P-(O'+d)| + |P-(O'-d)| = 2R .$$

If $d = c$ then for every $z$ in $E$ we have

$$|z-(O'+c)| + |z-(O'-c)| \leq 2R \text{ with equality for } z = P,P' .$$

Hence, $\frac{1}{n}\max_{z\in E}\left[\frac{n}{2}|z-(O'+d)| + \frac{n}{2}|z-(O'-d)|\right] = \frac{1}{2}(2R) = R$ .

Thus the set (3.2.16) for any $d \leq c = R\epsilon$ is a set of Chebyshev points of even order $n$ for $E$.

One can show that the set

$$z_i = \begin{cases} O' + d & \text{for } i = 1,2,\dots,\frac{n-1}{2} \\ O' & i = \frac{n+1}{2} \\ O' - d & i = \frac{n+3}{2},\dots,n \end{cases} \tag{3.2.17}$$

for any $d \leq c = R\epsilon$ is a set of Chebyshev points of odd order $n$ for $E$.

In the case where the Chebyshev points of order $n$ for $E$ are unique, the Chebyshev function of order $n$ for $E$ is also unique. It is

$$C_n^{(\lambda)}(z;E) = n^{\frac{1}{\lambda}}|z-O'| . \tag{3.2.18}$$

Since $m_n^{(\lambda)}(E) = n^{\frac{1}{\lambda}} R$

$$\mathcal{L}_{C_n^{(\lambda)}} : \{z \mid |z - 0'| \leq R\} = D_E \tag{3.2.19}$$

where $D_E$ is the <u>spanning disk of E</u>, i.e., the spanning circle and its interior.

$$\hat{\mathcal{L}}_{C_n^{(\lambda)}} : \{z \mid |z - 0'| = R\} = C_E \tag{3.2.20}$$

If E is a set such that Chebyshev points of even order n are not unique then

$$C_n^{(1)}(z;E) = \frac{n}{2}\left(|z - (0'+d)| + |z - (0'-d)|\right) \tag{3.2.21}$$

where $d \leq c$ . Hence

$$\mathcal{L}_{C_n^{(1)}} : \{z \mid |z - (0'+d)| + |z - (0'-d)| \leq 2R\} \tag{3.2.22}$$

and

$$\hat{\mathcal{L}}_{C_n^{(1)}} : \{z \mid |z - (0'+d)| + |z - (0'-d)| = 2R\} \tag{3.2.23}$$

$\mathcal{L}_{C_n^{(1)}}$ is an ellipse and its interior with foci at $(0'+d)$ and $(0'-d)$ and major axis of length 2R . It coincides with $D_E$ for $d = 0$ and with the ellipse $M_E$ and its interior for $d = c$ . In all cases the set $\hat{\mathcal{L}}_{C_n^{(1)}}$ is its boundary.

If $E$ is a set such that Chebyshev points of odd order $n$ are not unique, then

$$C_n^{(1)}(z;E) = \frac{n-1}{2}\left(|z-(O'+d)| + |z-(O'-d)|\right) + |z-O'| \tag{3.2.24}$$

where $d \le c$ . Hence

$$\mathscr{L}_{C_n^{(1)}}: \{z \,|\, \tfrac{n-1}{2}(|z-(O'+d)| + |z-(O'-d)|) + |z-O'| \le nR\} \tag{3.2.25}$$

and

$$\hat{\mathscr{L}}_{C_n^{(1)}}: \{z \,|\, \tfrac{n-1}{2}(|z-(O'+d)| + |z-(O'-d)|) + |z-O'| = nR\} \tag{3.2.26}$$

$\mathscr{L}_{C_n^{(1)}}$ is a superset of the ellipse with foci $O'+d$ and $O'-d$ and major axis of length $2R$ , and its interior.

Theorem 3.2.9

For any Case I set $E$ of diameter $d(E)$ and for $\lambda \ge 1$ , $\chi^{(\lambda)}(E) = \frac{d(E)}{2}$ .

For any Case II set $E$ of diameter $d(E)$ and for $\lambda \ge 1$ , $\frac{d(E)}{2} \le \chi^{(\lambda)}(E) \le \frac{d(E)\sqrt{3}}{3}$ .

Proof: Statements follow directly from Theorem 3.2.2 and Theorem 3.2.5.

The problem of calculating the transfinite diameter of a compact set in $R_2$ for $\lambda \ge 1$ is more difficult. Since $E \subset D_E$ , the monotony property of the transfinite diameter gives $\delta^{(\lambda)}(E) \le \delta^{(\lambda)}(D_E)$ . From the appendix, the results of Pólya and Szegö furnish us with the values of $\delta^{(\lambda)}(D_E)$ . Theorem 3.2.5 provides us with a lower bound for $\delta^{(\lambda)}(E)$ . Thus we have

$$R \leq \delta^{(\lambda)}_E \leq \frac{\Gamma(1+\lambda)^{\frac{1}{\lambda}}}{\Gamma\left(1+\frac{\lambda}{2}\right)^{\frac{2}{\lambda}}} R \quad \text{for} \quad 1 \leq \lambda \leq 2 \qquad (3.2.27)$$

$$R \leq \delta^{(\lambda)}(E) \leq (2^{1-\frac{1}{\lambda}}) R \quad \text{for} \quad \lambda \geq 2 \qquad (3.2.28)$$

<u>Theorem 3.2.10</u>

Let $E$ be a Case I set. Then $\delta^{(\lambda)}(E) = \left(2^{1-\frac{1}{\lambda}}\right) R$ for $\lambda \geq 2$ .

<u>Proof</u>: Since $E$ is a Case I set, $E \cap C_E \supset \{P,P'\}$ where $|P-P'| = 2R$ . Let $k$ be even and place $\frac{k}{2}$ points at $P$ and $\frac{k}{2}$ points at $P'$ . Then

$$\delta^{(\lambda)}_k(E) \geq \left[\frac{\left(\frac{k}{2}\right)\left(\frac{k}{2}\right)}{\binom{k}{2}} |P-P'|^{\lambda}\right]^{\frac{1}{\lambda}} = 2R \left[\frac{1}{2\left(1-\frac{1}{k}\right)}\right]^{\frac{1}{\lambda}}$$

for $\lambda > 0$ and as $k \to \infty$ we have

$$\delta^{(\lambda)}(E) \geq \left(2^{1-\frac{1}{\lambda}}\right) R \, . \qquad (3.2.29)$$

Since (3.2.28) gives us the reverse inequality we have $\delta^{(\lambda)}(E) = \left(2^{1-\frac{1}{\lambda}}\right)R$ for $\lambda \geq 2$ . In view of (3.2.1) we may also write this as

$$\delta^{(\lambda)}(E) = \frac{d(E)}{2^{\frac{1}{\lambda}}} \qquad (3.2.30)$$

for any Case I set $E$ of diameter $d(E)$ , for $\lambda \geq 2$ .

Theorem 3.2.11

Let $E$ be a Case I set. Then

$$\left(2^{1-\frac{1}{\lambda}}\right) R \leq \delta^{(\lambda)}(E) \leq \left[\frac{\Gamma(1+\lambda)^{\frac{1}{\lambda}}}{\Gamma(1+\frac{\lambda}{2})^{\frac{2}{\lambda}}}\right] R \qquad \text{for } 1 \leq \lambda \leq 2$$

Proof: The right hand side follows from (3.2.27) and the left hand side from (3.2.29).

Theorem 3.2.12

Let $E$ be a Case I set in $R_2$ . Then

1) For $\lambda > 1$ $\quad \delta^{(\lambda)}(E) > \chi^{(\lambda)}(E)$ .

2) For $\lambda = 1$ $\quad \delta^{(1)}(E) \geq \chi^{(1)}(E)$ with equality if and only if $E$ is "one dimensional," i.e., $E$ contains no points other than those on segment $PP'$ .

Proof:

1) From (3.2.29) we have $\delta^{(\lambda)}(E) \geq \left(2^{1-\frac{1}{\lambda}}\right) R$ for $\lambda \geq 1$ . Thus for $\lambda > 1$ $\delta^{(\lambda)}(E) > R = \chi^{(\lambda)}(E)$ .

2) The "if" part corresponds to Theorem 3.1.6. We shall show that if $E$ contains any point not lying on the segment $PP'$ then $\delta^{(1)}(E) > \chi^{(1)}(E)$ .

Suppose $Q \in E$ , $QP = QP'$ and $Q$ is not on segment $PP'$ . Thus $Q$ lies on the perpendicular bisector of $PP'$ but not on the segment itself. Let $PQ = A$ . We know $PP' = 2R$ .

Now if we place $k$ points on $P$ , $k$ on $P'$ , and $m$ on $Q$ where

$2k + m = n$ , we have

$$\begin{aligned} d_n^{(1)}(E) &\ge \frac{1}{\binom{n}{2}} [2km(PQ) + k^2(PP')] \\ &= \frac{1}{\binom{n}{2}} [2k(n-2k)A + 2k^2R] . \end{aligned} \qquad (3.2.31)$$

Let $f(k) = 2k(n-2k)A + 2k^2R$ . We note $f'(k) = (2nA - 8kA + 4kR)$ and $f''(k) = -8A + 4R < 0$ . (We are treating $k$ here as a real variable, and not restricting it to only integral values.) Hence $k = \frac{nA}{4A-2R}$ gives $f(k)$ its maximum value. Substituting in (3.2.31) and letting $n \to \infty$ gives

$$\delta^{(1)}(E) \ge \frac{A^2}{2A-R}$$

Now $A = R \sec \alpha$ where $\alpha$ is the angle between $PP'$ and $PQ$ . Since $Q$ does not lie on $PP'$ , $\alpha > 0$ . Hence $\sec \alpha > 1$ . Thus $(\sec \alpha - 1)^2 > 0$ or $\sec^2\alpha > 2 \sec \alpha - 1$ , or $\frac{A^2}{R^2} > 2\left(\frac{A}{R}\right) - 1$ or $\frac{A^2}{2A-R} > R$ . Thus $\delta^{(1)}(E) > R$ .

Now suppose there exists no such point $Q$ . Let $Q'$ be any other point of $E$ which does not lie on segment $PP'$ . Either $PQ'$ or $P'Q'$ intersects the perpendicular bisector of $PP'$ . Suppose $PQ'$ intersects it. Let $Q$ be the point of intersection. Then $PQ' = PQ + QQ'$ . Suppose we place $k$ points at $P$ , $k$ points at $P'$ and $m$ points at $Q'$ where $2k + m = n$ . Then

$$\begin{aligned} d_n^{(1)}(E) &\geq \frac{1}{\binom{n}{2}}[km(PQ' + P'Q') + k^2PP'] \\ &= \frac{1}{\binom{n}{2}}[km(PQ + QQ' + P'Q') + k^2PP'] \\ &> \frac{1}{\binom{n}{2}}[km(PQ + P'Q) + k^2PP'] \\ &= \frac{1}{\binom{n}{2}}[2km(PQ) + k^2PP'] \end{aligned}$$

The last expression is the same as (3.2.31). But we have shown for an expression of this form $\delta^{(1)}(E) > R = \chi^{(1)}(E)$ .

Let $\mathcal{T}$ be the class of all acute triangles $A_1A_2A_3$ such that $\sin\alpha_1 + \sin\alpha_2 + \sin\alpha_3 > \frac{9}{4}$ . Let $\mathcal{T}'$ be the class of all acute isosceles triangles.

<u>Theorem 3.2.13</u>

Let $E$ be a Case II set. If $E \cap C_E \supset \{A_1, A_2, A_3\}$ where $\Delta A_1A_2A_3 \in \mathcal{T}$ or $\Delta A_1A_2A_3 \in \mathcal{T}'$ , then $\delta^{(\lambda)}(E) > \chi^{(\lambda)}(E)$ for $\lambda \geq 1$ .

<u>Proof:</u> Suppose $E \cap C_E \supset \{A_1A_2A_3\}$ and $\Delta A_1A_2A_3 \in \mathcal{T}$ . Let $n = 3k$ . Place $k$ points at $A_1$ , $k$ at $A_2$ , and $k$ at $A_3$ . Then for $\lambda = 1$ we have

$$d_n^{(1)}(E) = \max \frac{1}{\binom{n}{2}} \sum_{1\leq i<j\leq n} |x_i - x_j| \geq \frac{1}{\binom{n}{2}} \frac{n^2}{9}\left[A_1A_2 + A_2A_3 + A_3A_1\right]$$

$$= \frac{2}{9\left(1 - \frac{1}{n}\right)}\left[A_1A_2 + A_2A_3 + A_3A_1\right]$$ and as $n \to \infty$ , we have

$\delta^{(1)}(E) \geq \frac{2}{9}\left[A_1A_2 + A_2A_3 + A_3A_1\right]$ . One can show that if $R$ is the radius of the circumscribed circle of acute triangle $A_1A_2A_3$ , then

$R = \frac{1}{2}\left[\frac{A_1A_2 + A_2A_3 + A_3A_1}{\sin\alpha_3 + \sin\alpha_1 + \sin\alpha_2}\right]$ where $\alpha_i$ is the angle at the vertex $A_i$ for $i = 1,2,3$. Thus $\delta^{(1)}(E) \geq \frac{4}{9} R\,(\sin\alpha_1 + \sin\alpha_2 + \sin\alpha_3) > R = \chi^{(1)}(E)$. For $\lambda > 1$, $\delta^{(\lambda)}(E) \geq \delta^{(1)}(E) > R = \chi^{(\lambda)}(E)$.

Suppose $E \cap C_E \supset \{A_1A_2A_3\}$ and $\Delta\, A_1A_2A_3 \in \mathcal{T}'$ ; Let $A_1A_3 = A_2A_3$ and place $k$ points at $A_1$, $k$ at $A_2$, and $m$ at $A_3$ where $2k + m = n$. Following the same argument as in the proof of Theorem 3.2.12 gives us a function

$$d_n^{(1)}(E) \geq \frac{1}{\binom{n}{2}}\left[2k(n-2k)A_1A_3 + k^2\, A_1A_2\right] \qquad (3.2.32)$$

As a continuous function of $k$, $d_n^{(1)}(E)$ reaches a maximum at $k = \frac{nA_1A_3}{4A_1A_3 - A_1A_2}$. Substituting this in (3.2.32) and letting $n \to \infty$ gives $\delta^{(1)}(E) \geq \frac{2(A_1A_3)^2}{4A_1A_3 - A_1A_2}$. Now $2R = \frac{A_1A_3}{\sin\alpha_1} = \frac{A_1A_2}{\sin\alpha_3}$. This gives $\delta^{(1)}(E) \geq R\left[\frac{2\sin\alpha_1}{2-\cos\alpha_1}\right]$ but since $\Delta\, A_1A_2A_3$ is acute, $\alpha_1 > \frac{\pi}{4}$. Hence $0 < \cos\alpha_1 < \frac{\sqrt{2}}{2}$. Thus $\cos\alpha_1 < \frac{4}{5}$ or $5\cos^2\alpha_1 < 4\cos\alpha_1$. Now $4\cos^2\alpha_1 < 4\cos\alpha_1 - \cos^2\alpha_1$ or $4 - 4\cos^2\alpha_1 > \cos^2\alpha_1 - 4\cos\alpha_1 + 4$. Hence $(2\sqrt{1-\cos^2\alpha_1})^2 > (2-\cos\alpha_1)^2$ or $2\sin\alpha_1 > 2-\cos\alpha_1$. Therefore $\delta^{(1)}(E) > R = \chi^{(1)}(E)$ and $\delta^{(\lambda)}(E) \geq \delta^{(1)}(E) > R = \chi^{(\lambda)}(E)$.

## 3.3 n-dimensional Real Euclidean Space

Our aim in this section will be to establish the result $\chi^{(\lambda)}(E) = R$ for $\lambda \geq 1$ where $E$ is any compact set in $R_n$ and $R$

is the radius of its spanning sphere. We shall follow along the lines developed in Section 3.2.

Definition 3.3.1

Let $E$ be a compact set in $R_n$. We shall denote by $S_E$ the sphere of largest dimension $S_{n-1} : \sum_{i=1}^{n} (x_i - h_i)^2 = R^2$ of smallest radius $R$ which encloses the set $E$. $S_E$ shall be called the span-ning sphere of $E$.

The proof of Theorem 3.2.1 printed in Yaglom [33] is purely geometric, greatly relying on one's 2-dimensional visualization. Thus we must develop analytic methods to extend its result to $n$ dimensions. This is done in the series of lemmas below.

Lemma 3.3.1

The spanning sphere of $E$ is unique.

Proof: Suppose not. Let $S_E : \left\{(x_1,x_2,\ldots,x_n) \in R_n \mid \sum_{i=1}^{n} (x_i - h_i)^2 = R^2\right\}$ and $S'_E : \left\{(x_1,x_2,\ldots,x_n) \in R_n \mid \sum_{i=1}^{n} (x_i - g_i)^2 = R^2\right\}$ both be spanning spheres of $E$. Let $B_E : \left\{(x_1,x_2,\ldots,x_n) \in R_n \mid \sum_{i=1}^{n} (x_i - h_i)^2 \le R^2\right\}$ and $B'_E : \left\{(x_1,x_2,\ldots,x_n) \in R_n \mid \sum_{i=1}^{n} (x_i - g_i)^2 \le R^2\right\}$. Then since $E \subset B_E$ and $E \subset B'_E$, we have $E \subset B_E \cap B'_E = F \neq \emptyset$. Hence $\sum_{i=1}^{n} (h_i - g_i)^2 < 4R^2$.

Now a hyperplane through the intersection of $S_E$ and $S'_E$ takes the form:

$$\pi : \left\{(x_1,x_2,\dots,x_n) \in R_n \,\middle|\, (H - G,X) = b\right\} \qquad (3.3.1)$$

where $H = (h_1,h_2,\dots,h_n)$ , $G = (g_1,g_2,\dots,g_n)$ and $b = \frac{(H,H) - (G,G)}{2}$ . We note that the line $\ell$ through $H$ and $G$ is perpendicular to the hyperplane. Let $\ell \cap \pi = M = (m_1,m_2,\dots,m_n)$ . Let the positive direction of line $\ell$ be the direction from $H$ to $G$ .

Now $\pi$ is a separating hyperplane between $H$ and $G$ . For if

$$(H\text{-}G,H) > b \quad \text{and} \quad (H\text{-}G,G) > b \qquad (3.3.2)$$

then we could write

$$H = \bar{t}_1 G + (1 - \bar{t}_1)M \quad \text{for} \quad 0 < \bar{t}_1 < 1 \ . \qquad (3.3.3)$$

But then $(H\text{-}G,H\text{-}M) = (H\text{-}G,H) - (H\text{-}G,M) = (H\text{-}G,H) - b > 0$ from (3.3.2) and $(H\text{-}G,H\text{-}M) = ((\bar{t}_1 - 1)G\text{-}M, \bar{t}_1(G\text{-}M)) = \bar{t}_1(\bar{t}_1\text{-}1)(G\text{-}M,G\text{-}M)$ from (3.3.3). Since $0 < \bar{t}_1 < 1$ this implies $(H\text{-}G,H\text{-}M) < 0$ . Contradiction! If we assume

$$(H\text{-}G,H) < b \quad \text{and} \quad (H\text{-}G,G) < b \qquad (3.3.4)$$

then we could write

$$G = \bar{t}_2 H + (1 - \bar{t}_2)M \quad \text{for} \quad 0 < \bar{t}_2 < 1 \ . \qquad (3.3.5)$$

Now $(H\text{-}G,G\text{-}M) = (H\text{-}G,G) - (H\text{-}G,M) = (H\text{-}G,G) - b < 0$ from (3.3.4). But $(H\text{-}G,G\text{-}M) + ((1\text{-}\bar{t}_2)H\text{-}M,\bar{t}_2(H\text{-}M)) = (1\text{-}\bar{t}_2)\bar{t}_2(H\text{-}M,H\text{-}M)$ from (3.3.5). Since $0 < \bar{t}_2 < 1$ we have $(H\text{-}G,G\text{-}M) > 0$ . Contradiction! Thus $(H\text{-}G,X) = b$ is a separating hyperplane. Let

$$F_1 = \left\{(y_1,y_2,\ldots,y_n) \in F \,\middle|\, (H-G,Y) > b\right\}$$

$$F_2 = \left\{(y_1,y_2,\ldots,y_n) \in F \,\middle|\, H-G,Y) \le b\right\}$$

Let $Y = (y_1,y_2,\ldots,y_n) \in F_1$ . We have $R^2 = \sum_{i=1}^{n} (y_i - g_i)^2 =$ $\sum_{i=1}^{n} (y_i - m_i + m_i - g_i)^2 = \sum_{i=1}^{n} (y_i - m_i)^2 + \sum_{i=1}^{n} (m_i - g_i)^2 +$ $2\sum_{i=1}^{n} (y_i - m_i)(m_i - g_i)$ . Now since $M$ lies between $H$ and $G$ on line $\ell$ there exists a $t$ , $0 < t < 1$ , such that $M - G = t(H-G)$ . Thus $\sum_{i=1}^{n} (y_i - m_i)(m_i - g_i) = (Y-M,M-G) = (Y-M,t(H-G)) = t[(Y,H-G)-(M,H-G)]$ . Since $Y \in F_1$ , $(Y,H-G) > b$ . Hence $(Y-M,M-G) > 0$ . Thus $\sum_{i=1}^{n} (y_i - m_i)^2 = R^2 - (M-G,M-G) - 2(Y-M,M-G) < R^2$ . Now let $Y = (y_1,y_2,\ldots,y_n) \in F_2$ . We have $R^2 = \sum_{i=1}^{n} (y_i - h_i)^2 =$

$$\sum_{i=1}^{n} (y_i - m_i + m_i - h_i)^2 = \sum_{i=1}^{n} (y_i - m_i)^2 + \sum_{i=1}^{n} (m_i - h_i)^2 + 2\sum_{i=1}^{n} (y_i - m_i)(m_i - h_i)$$

Since $M$ lies between $H$ and $G$ , there exists a $t$ , $-1 < t < 0$ such that $M-H = t(H-G)$ . Thus

$$\sum_{i=1}^{n} (y_i - m_i)(m_i - h_i) = (Y-M,M-H) = (Y-M,t(H-G)) = t[(Y,H-G) - (M,H-G)] .$$

Since $Y \in F_2$ , $(Y,H-G) \le b$ . Hence $(Y-M,M-H) \ge 0$ . Thus

$$\sum_{i=1}^{n} (y_i - m_i)^2 = R^2 - (M-H,M-H) - 2(Y-M,M-H) < R^2 .$$

Let $R'^2 = \max[R^2 - (M-H,M-H) , R^2 - (M-G,M-G)]$ . Consider $\left\{(y_1,y_2,\ldots,y_n) \in R_n \,\middle|\, \sum_{i=1}^{n} (y_i - m_i)^2 \le R'^2\right\}$ . All $Y \in F$ belong to this set. Hence all $Y \in E$ belong to this set.

Hence $\sum_{i=1}^{n} (y_1 - m_i)^2 = R'^2$ is a spanning sphere for $E$ .

But $R' < R$ . Contradiction! Thus the spanning sphere of $E$ is unique.

Lemma 3.3.2

Let $E$ be compact in $R_n$ and let $S_E$ be the spanning sphere of $E$ . Then $E \cap S_E \neq \emptyset$ .

Proof: Suppose $E \cap S_E = \emptyset$ . Let $\rho(E,S_E) = \mu$ . Consider a sphere $S_E'$ with same center as $S_E$ and with radius $R - \mu$ where $R$ is the radius of $S_E$ . Then $E$ is enclosed by $S_E'$ , but $S_E'$ has smaller radius than $S_E$ . Contradiction! Hence $E \cap S_E \neq \emptyset$ .

Lemma 3.3.3

Let $S$ be a sphere which encloses $E$ with center $H = (h_1,h_2,\ldots,h_n)$ and radius $R$ . Let $Q = S \cap E$ . Suppose $|Q| = k$ . Then if $H$ is not in the convex hull of $Q$ , the elements of $Q$ lie in an open hemisphere of $S$ .

Proof: We note first that $k$ need not be finite. Let $V$ be the convex hull of $Q$ . $V$ is closed since $Q$ is compact. Since $H \notin V$ , there exists a hyperplane $\pi$ which strictly separates $H$ and $V$ . Let $\pi : \{(x_1,x_2,\ldots,x_n) \in R_n | (P,X) = b\}$ . Then

$$(P,H) = a > b \quad \text{and} \quad (P,Y) < b \tag{3.3.6}$$

for all $Y \in V$ and hence for all $Y \in Q$ .

Consider the set $\{(x_1,x_2,\ldots,x_n) \in R_n | (X-H,X-H) = R^2 \text{ and } (P,X) < a\}$ . This set is an open hemisphere and we claim all $Y \in Q$ are in it. Since $Y \in Q = E \cap S$ , $Y \in S$ . Hence $(Y-H,Y-H) = R^2$ and from (3.3.6) $(P,Y) < b < a$ .

Lemma 3.3.4

Let $S$ be a sphere which encloses $E$ and let $Q = S \cap E$ . If the elements of $Q$ lie in an open hemisphere of $S$ , then $S$ is not the spanning sphere of $E$ .

Proof: Without loss of generality we shall write $S : \{(x_1, x_2, \ldots, x_n) \in R_n \mid \sum_{i=1}^{n} x_i^2 = R^2\}$ . Suppose the elements of $Q$ lie in some open hemisphere of $S$ , say the hemisphere designated by the set $\{(x_1, x_2, \ldots, x_n) \in R_n \mid (X,X) = R^2$ and $(P,X) > 0\}$ .

The vector $P$ is perpendicular to the hyperplane $(P,X) = 0$ . Suppose we move $S$ a distance $v$ along the vector $P$ . We shall choose $v$ such that $E$ is enclosed by $S'$ and $E \cap S' = \emptyset$ , where $S'$ is the translate of $S$ . Then by Lemma 3.3.2, $S'$ is not the spanning sphere of $E$ . Since the translation is a rigid motion, the radius of $S'$ is equal to the radius of $S$ . Hence $S$ is not the spanning sphere of $E$ , and we shall be finished.

Let $b$ be the distance from the set $Q$ to the hyperplane $(P,X) = 0$ . Choose any integer $m > 1$ such that

$$b > \frac{R}{m} . \tag{3.3.7}$$

We consider the set $E$ in three parts. Let

$$\begin{aligned} E_1 &= \left\{(x_1, x_2, \ldots, x_n) \in E \mid (P,X) \le 0\right\} \\ E_2 &= \left\{(x_1, x_2, \ldots, x_n) \in E \mid 0 \le (P,X) \le \frac{R}{m}\right\} \\ E_3 &= \left\{(x_1, x_2, \ldots, x_n) \in E \mid (P,X) > \frac{R}{m}\right\} \end{aligned} \tag{3.3.8}$$

Let

$$F_1 = \left\{(x_1,x_2,\ldots,x_n) \in R_n \,\middle|\, (X,X) = R^2 \text{ and } (P,X) \le 0\right\}$$
$$F_2 = \left\{(x_1,x_2,\ldots,x_n) \in R_n \,\middle|\, (X,X) = R^2 \text{ and } 0 \le (P,X) \le \frac{R}{m}\right\} \qquad (3.3.9)$$

Let $\delta_1$ be the distance from $E_1$ to $F_1$ and $\delta_2$ be the distance from $E_2$ to $F_2$ . Finally we let

$$c = \min\left(\delta_1 , \delta_2 , \frac{2R}{\|P\| m}\right) \qquad (3.3.10)$$

We note that $c \neq 0$ since all points of contact between E and S belong to $E_3$ . Take a fixed $\epsilon$ , $0 < \epsilon < 1$ and let $\mu > 0$ be such that $\mu^2(P,P) = \epsilon^2 c^2$ . Thus the center of our translated sphere $S'$ is $(\mu p_1, \mu p_2, \ldots, \mu p_n)$ and our translated distance is $v = \epsilon c$ . The square of the distance from any point $\bar{X} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ to the center of $S'$ is given by

$$d^2 = \sum_{i=1}^{n} (\bar{x}_i - \mu p_i)^2 = (\bar{X},\bar{X}) - 2\mu(\bar{X},P) + \mu^2(P,P) \qquad (3.3.11)$$

Let $\bar{X} \in E_1$ . Then $(\bar{X},\bar{X}) \le (R - \delta_1)^2 \le (R - c)^2$ . From Schwartz' inequality we have $|(\bar{X},P)| \le \|P\|\,\|\bar{X}\|$ . Hence $(\bar{X},P) \ge -\|P\|\,\|X\|$ . Thus $-2\mu(\bar{X},P) \le 2\mu\|P\|(R - c) = 2\epsilon c(R - c)$ . Thus $d^2 \le (R - c)^2 + 2\epsilon c(R - c) + \epsilon^2 c^2 = [(R - c) + \epsilon c]^2$ , or $d \le R + (\epsilon - 1)c$ . But $\epsilon < 1$ and $c \neq 0$ . Thus $d < R$ .

Let $\bar{X} \in E_2$ . Then $(\bar{X},\bar{X}) \le (R - \delta_2)^2 \le (R - c)^2$ and $(P,\bar{X}) \ge 0$ . Hence $d^2 \le (R - c)^2 + \epsilon^2 c^2 < (R - c)^2 + c^2 < R^2 - c^2 + c^2 = R^2$ , or $d < R$ .

Let $\bar{X} \in E_3$ . Then $(\bar{X},\bar{X}) \le R^2$ and $(P,\bar{X}) > \frac{R}{m}$ . Thus $d^2 < R^2 - 2\mu \frac{R}{m} + \mu^2(P,P) = R^2 - \frac{2\epsilon cR}{\|P\| m} + \epsilon^2 c^2$ . But $c \le \frac{2R}{\|P\| m}$ . Thus

$ec \le \frac{2\epsilon R}{\|P\| m}$ and $\epsilon^2 c^2 \le \frac{2\epsilon^2 cR}{\|P\| m} < \frac{2\epsilon cR}{\|P\| m}$ . Thus $d < R$.. So we have shown that $E \cap S' = \emptyset$, and by the previous remarks the lemma is proved.

We are finally led to the following analogue of Theorem 3.2.1.

Theorem 3.3.1

Let $E$ be a compact set in $R_n$. The spanning sphere $S_E$ is unique and its center $O'$ lies in the convex hull of $Q = S_E \cap E$. Moreover, $|Q| = k \geq 2$.

Proof: From Lemma 3.3.3, if $O'$ does not lie in the convex hull of $Q$, then the elements of $Q$ lie in an open hemisphere of $S_E$. But from Lemma 3.3.3, this implies $S_E$ is not the spanning sphere of $E$. Contradiction! Thus $O'$ is in the convex hull of $Q$. We have the cardinality of $Q$ greater than or equal to $2$ since $O'$ cannot be in the convex hull of less than two points. Uniqueness of $S_E$ is the result of Lemma 3.3.1.

If the cardinality of $Q$ is $k$, and $k$ is finite, we shall see that the development of the main theorem in $R_n$ proceeds much as it does in $R_2$. However, suppose the cardinality of $Q$ is infinite? Shall we be forced to consider infinite sums? Fortunately, no. A theorem of Cartheodory is available.

Theorem (Cartheodory)

Let $Q$ be a subset of $R_n$ and $O'$ a point in the convex hull of $Q$. Then there exists a subset $Q_1$ of $Q$, containing at most $n+1$ points such that $O'$ is in the convex hull of $Q_1$.

Proof: Refer to Eggleston, H.G., Convexity.

Thus in the following material when we write $\{A_1, A_2, \ldots, A_k\}$ we shall mean the elements of Q if the cardinality of Q is finite, or we shall mean the subset of Q of at most n+1 points which contain the center of $S_E$ in its convex hull, if the cardinality of Q is infinite.

The reverse λ-Steiner problem in $R_n$ takes the following form: If $A_1, A_2, \ldots, A_k$ $k \geq 2$ are k points lying on a sphere $S_{n-1}$ in $R_n$ such that they do not all lie in an open hemisphere of $S_{n-1}$, can we attach non-negative weights, $w_{A_1}, w_{A_2}, \ldots, w_{A_k}$ to the points $A_1, A_2, \ldots, A_k$ so as to make any given point **Q** in the convex hull of $\{A_1, A_2, \ldots, A_k\}$ a λ-Steiner point? i.e., to minimize a function of the form

$$\sum_{i=1}^{k} w_{A_i} \|A_i - Q\|^{\lambda} \qquad (3.3.12)$$

for $\lambda \geq 1$, where $\|A_i - Q\|$ denotes the Euclidean distance from the point $A_i$ to Q ? We shall show that the answer to the question is affirmative. Again we shall use approximation theory.

Let E be a compact set in $R_n$ and let μ be a positive measure on $R_n$. Consider the space of continuous functions whose domain is

E and whose range lies in some inner product space H . This space, which we denote by C(E;H) can be made into a normed linear space in various ways. We shall define a norm thusly: If $f \in C(E;H)$ then by $|||f|||_p$ , $p \geq 1$ , we shall mean $\left(\int_E \|f(x)\|^p \, d\mu\right)^{\frac{1}{p}}$ where $\|f(x)\|$ denotes the inner-product space norm of the range value f(x) .

If V is a finite-dimensional subspace of C(E;H) we may pose an approximation problem by asking: Given $f \in C(E;H)$ . What is a best approximation Q to f in p-norm out of V ? i.e., find a $Q \in V$ such that $|||f - X|||_p \geq |||f - Q|||_p$ or equivalently

$$\int_E \|f(x) - X(x)\|^p \, d\mu \geq \int_E \|f(x) - Q(x)\|^p \, d\mu \qquad (3.3.13)$$

for all $X \in V$ .

This approximation problem has a solution characterized by Theorem 3.3.2. But first we must extend the definition of sgn to vectors.

<u>Definition 3.3.2</u>

If $s \in R_n$ , we define $\operatorname{sgn} s = \frac{s}{\|s\|}$ for $s \neq 0$ and $\operatorname{sgn} s = 0$ for $s = 0$ .

<u>Theorem 3.3.2</u>

For $p \geq 1$ , if $\int_E \|f(x) - Q(x)\|^{p-1} \left(X(x), \operatorname{sgn}(f(x) - Q(x))\right) d\mu = 0$ for all $X \in V$ , then Q is a best approximation to f in p-norm out of V .

<u>Proof:</u>

$$\int_E \|f(x)-Q(x)\|^p \, d\mu = \int_E \|f(x)-Q(x)\|^{p-1} \left(f(x)-Q(x), \operatorname{sgn}(f(x)-Q(x))\right) d\mu$$

$$= \int_E \|f(x)-Q(x)\|^{p-1} \Big(f(x)-X(x)\,,\ \operatorname{sgn}(f(x)-Q(x))\Big)\,d\mu$$

$$+ \int_E \|f(x)-Q(x)\|^{p-1} \Big(X(x)-Q(x)\,,\ \operatorname{sgn}(f(x)-Q(x))\Big)\,d\mu$$

The second term on the right is zero by hypothesis. From the Schwarz inequality we have

$$\int_E \|f(x) - Q(x)\|^p \,d\mu \le \int \|f(x) - Q(x)\|^{p-1} \,\|f(x) - X(x)\|\,d\mu$$

If $p = 1$ , the inequality (3.3.13) results. If $p > 1$ we apply Holder's inequality to the integral on the right. Thus

$$\int_E \|f(x) - Q(x)\|^p \,d\mu \le \Big(\int_E \|f(x) - Q(x)\|^{(p-1)q}\Big)^{\frac{1}{q}} \Big(\int_E \|f(x) - X(x)\|^p \,d\mu\Big)^{\frac{1}{p}}$$

or

$$\Big(\int_E \|f(x) - Q(x)\|^p \,d\mu\Big)^{\frac{1}{p}} \le \Big(\int_E \|f(x) - X(x)\|^p \,d\mu\Big)^{\frac{1}{p}}$$

Raising both sides to the p-th power we have (3.3.13).

Now we take $E$ to be the finite set of points $\{A_1, A_2, \ldots, A_k\}$ and choose the measure $\mu$ to assign the weights $w_{A_i}$ at $A_i$ , $i = 1,2,\ldots,k$ . Let $H$ be k-dimensional complex space with the inner-product defined by: If $z = (z_1, z_2, \ldots, z_k) \in H$ and $w = (w_1, w_2, \ldots, w_k) \in H$ , then, $(z,w) = \sum_{i=1}^{k} z_i \bar{w}_i$ . Let $f$ be the identity function on $E$ and let the subspace $V$ consist of the constant functions. Then, setting $p = \lambda$ , the right hand side of (3.3.13) becomes $\sum_{i=1}^{k} \|A_i - Q\|^{\lambda} w_{A_i}$ which is identical to the function (3.3.12). The sufficiency condition of Theorem 3.3.2 becomes

$$\sum_{i=1}^{k} w_{A_i} \|A_i - Q\|^{p-1} (X, \operatorname{sgn}(A_i - Q)) = 0 \qquad (3.3.14)$$

for all constants $X$ .

Theorem 3.3.3

$w_{A_i} > 0$ may be chosen so as to satisfy the sufficiency condition (3.3.14).

Proof: See Theorem 3.2.4. We set

$$w_{A_i} = \lambda_i \|A_i - Q\|^{2-p} \quad \text{for} \quad i = 1,2,\ldots,k \qquad (3.3.15)$$

where the $\lambda_i$'s are the barycentric coordinates of $Q$ relative to the points $A_1, A_2, \ldots, A_k$ .

Thus all machinery has been developed to enable us to prove the n-dimensional analogue of the main theorem, Theorem 3.2.5.

Theorem 3.2.4

Let $E$ be a compact set in $R_n$ . Let $S_E$ be the spanning sphere of $E$ . Then $\chi^{(\lambda)}(E) = R$ for $\lambda \geq 1$ , where $R$ is the radius of $S_E$ .

Proof: Let $O'$ be the center of $S_E$ . Placing $X = O'$ for $i = 1,2,\ldots,m$ gives

$$\tau_m^{(\lambda)}(E) \leq \max_{X \in E} \left[ \frac{1}{m} \sum_{i=1}^{m} \|X - X_i\|^{\lambda} \right]^{\frac{1}{\lambda}} = \left( \frac{1}{m} m R^{\lambda} \right)^{\frac{1}{\lambda}} = R$$

Now suppose $X_i^*$, $i = 1,2,\ldots,m$ is a set of Chebyshev points of order $m$ for $E$ . We know $E \cap S_E \supset \{A_1, A_2, \ldots, A_k\}$ , $k \geq 2$ .

Now

$$T_m^{(\lambda)}(E) = \max_{X\in E} \left(\frac{1}{m}\sum_{i=1}^{m} \|X - X_i^*\|^\lambda\right)^{\frac{1}{\lambda}} \geq \max_{X\in\{A_1,A_2,\dots,A_k\}} \left(\frac{1}{m}\sum_{i=1}^{m} \|X - X_i^*\|^\lambda\right)^{\frac{1}{\lambda}}$$

$$\geq \left[\frac{w_{A_1}\left(\frac{1}{m}\sum_{i=1}^{m}\|A_1-X_i^*\|^\lambda\right) + w_{A_2}\left(\frac{1}{m}\sum_{i=1}^{m}\|A_2-X_i^*\|^\lambda\right) + \dots + w_{A_k}\left(\frac{1}{m}\sum_{i=1}^{m}\|A_k-X_i^*\|^\lambda\right)}{\sum_{j=1}^{k} w_{A_j}}\right]^{\frac{1}{\lambda}}$$

$$= \left[\frac{\frac{1}{m}\sum_{j=1}^{k} w_{A_j}\|A_j-X_1^*\|^\lambda + \frac{1}{m}\sum_{j=1}^{k} w_{A_j}\|A_j-X_2^*\|^\lambda + \dots + \frac{1}{m}\sum_{j=1}^{k} w_{A_j}\|A_j-X_m^*\|^\lambda}{\sum_{j=1}^{k} w_{A_j}}\right]^{\frac{1}{\lambda}} \qquad (3.3.16)$$

We choose $w_{A_j}$, $j = 1,2,\dots,k$ such that the point $O'$ is the $\lambda$-Steiner point, as in accordance with (3.3.15) with $Q = O'$. This can be done since we have shown that $O'$ is in the convex hull of $\{A_1, A_2,\dots,A_k\}$, $k \geq 2$. The last expression (3.3.16) thus is greater than or equal to

$$\left[\frac{\frac{1}{m}\sum_{j=1}^{k} w_{A_j}\|A_j-O'\|^\lambda + \frac{1}{m}\sum_{j=1}^{k} w_{A_j}\|A_j-O'\|^\lambda + \dots + \frac{1}{m}\sum_{j=1}^{k} w_{A_j}\|A_j-O'\|^\lambda}{\sum_{j=1}^{k} w_{A_j}}\right]^{\frac{1}{\lambda}}$$

$= R$ . Hence $T_m^{(\lambda)}(E) = R$ for every $m$ and $\chi^{(\lambda)}(E) = R$ .

As a final word we should mention the bounds for $R$ given by the n-dimensional version of Jung's theorem.

Theorem 3.3.5

Let $E$ be a compact set in $R_n$ with diameter $d(E)$ . Then its spanning sphere $S_E$ has a radius $R$ where

$$\frac{d(E)}{2} \leq R \leq \sqrt{\frac{n}{2(n+1)}}\, d(E) .$$

Proof: See Jung [17].

# CHAPTER 4

## Related Topics

In Chapter 2 we noted that one could extend the classical definitions of transfinite diameter and Chebyshev constant to more general metric spaces. We find that for spaces that are "centered", we can calculate the Chebyshev constant for any bounded subset for $\lambda \geq 1$ . The idea of "centeredness" is introduced in an entirely different context in a paper by Kolmogorov and Tihomirov [19], but it plays an important role in our theory.

We shall then obtain bounds for the Chebyshev constant of unit balls in normed linear spaces which hold for all averaging processes, and compare our results with some work of Hille dealing with the transfinite diameter of such sets.

### 4.1 Centered Spaces

#### Definition 4.1.1

Let $(X,\rho)$ be a metric space. A bounded subset $V \subseteq X$ of topological diameter $d(V)$ is said to be a centered set if there exists a point $x_V \in X$ such that $\rho(v,x_V) \leq \frac{d(V)}{2}$ for all $v \in V$ . A point $x_V$ with this property will be called a center of $V$ .

We note that a center of $V$ does not have to lie in $V$ . Centered subsets of metric spaces are of interest since one can find their Chebyshev constant for $\lambda \geq 1$ by virtue of the following theorem:

#### Theorem 4.1.1

Let $(X,\rho)$ be a metric space. If $W$ is a centered subset of $X$ of diameter $d(W)$ , then $\chi^{(\lambda)}(W) = \frac{d(W)}{2}$ for $\lambda \geq 1$ .

Proof:

$$\tau_n^{(\lambda)}(W) = \inf_{\substack{x_i \in X \\ i=1,2,\dots,n}} \sup_{w \in W} \left[ \frac{1}{n} \sum_{i=1}^{n} \rho(w,x_i)^{\lambda} \right]^{\frac{1}{\lambda}}$$

. Let $x_W \in X$ be a center of $W$ . Set $x_i = x_W$ for $i = 1,2,\dots,n$ . Then

$$\tau_n^{(\lambda)}(W) \leq \sup_{w \in W} \left[ \frac{1}{n}\, n\rho(w,x_W)^{\lambda} \right]^{\frac{1}{\lambda}} = \sup_{w \in W} \rho(w,x_W) \leq \frac{d(W)}{2} . \qquad (4.1.1)$$

If $W$ contains only one point, both $d(W)$ and $\chi^{(\lambda)}(W)$ are zero and the theorem is proved. Hence we suppose $W$ contains more than one point. Let $\bar{w}$ and $\underline{w}$ be two distinct elements of $W$ . For every $\epsilon > 0$ , there exists $\bar{x}_1,\bar{x}_2,\dots,\bar{x}_n \in X$ such that

$$\sup_{w \in W} \left[ \frac{1}{n} \sum_{i=1}^{n} \rho(w,\bar{x}_i)^{\lambda} \right]^{\frac{1}{\lambda}} - \epsilon < \tau_n^{(\lambda)}(W) . \qquad (4.1.2)$$

The triangle inequality gives

$$\rho(\underline{w},\bar{w}) \leq \rho(\underline{w},\bar{x}_i) + \rho(\bar{x}_i,\bar{w}) \quad \text{for} \quad i = 1,2,\dots,n .$$

Hence

$$n\left[\rho(\underline{w},\bar{w})\right]^{\lambda} \leq \sum_{i=1}^{n} \left[\rho(\underline{w},\bar{x}_i) + \rho(\bar{x}_i,\bar{w})\right]^{\lambda}$$

. Raising to the $\frac{1}{\lambda}$-th power, using Minkowski's inequality, and dividing through gives

$$\rho(\underline{w},\bar{w}) \leq \left[\frac{1}{n} \sum_{i=1}^{n} \rho(\underline{w},\bar{x}_i)^{\lambda}\right]^{\frac{1}{\lambda}} + \left[\frac{1}{n} \sum_{i=1}^{n} \rho(\bar{w},\bar{x}_i)^{\lambda}\right]^{\frac{1}{\lambda}}$$

. Now, taking the supremum over all pairs $\underline{w},\bar{w}$ in $W$ we have $d(W) \leq 2 \sup_{w \in W} \left[\frac{1}{n} \sum_{i=1}^{n} \rho(w,\bar{x}_i)^{\lambda}\right]^{\frac{1}{\lambda}}$ . From (4.1.2) this gives $\frac{d(W)}{2} < \tau_n^{(\lambda)}(W) + \epsilon$ . Since $\epsilon$ was arbitrary we have $\frac{d(W)}{2} \leq \tau_n^{(\lambda)}(W)$ and with (4.1.1) as $n \to \infty$ , $\chi^{(\lambda)}(W) = \frac{d(W)}{2}$ .

Thus for any ball $S$ in a metric space, Theorem 4.1.1 gives

$$\chi^{(\lambda)}(S) = r \tag{4.1.3}$$

where $r$ is the radius of $S$ and $\lambda \geq 1$. If we designate by $U$ the unit ball of a normed linear space, i.e., the set of all elements of norm less than or equal to one, then in particular we have

Corollary 4.1.1

For any normed linear space $X$ with unit ball $U$, $\chi^{(\lambda)}(U) = 1$ for $\lambda \geq 1$.

Definition 4.1.2

If every bounded subset of a metric space $(X,\rho)$ is a centered set, then we shall say $X$ is a centered space.

Corollary 4.1.2

Let $(X,\rho)$ be a centered metric space. If $V$ is any bounded subset of $X$ of diameter $d(V)$, then $\chi^{(\lambda)}(V) = \frac{d(V)}{2}$ for $\lambda \geq 1$.

An analysis of the proof of Theorem 4.1.1 (or Corollary 4.1.2) shows that it is the centeredness of the set $V$ (or the space $X$) which establishes the inequality $\chi^{(\lambda)}(V) \leq \frac{d(V)}{2}$. The reverse inequality holds for all bounded subsets of metric spaces, whether centered or not, by virtue of Minkowski's inequality. Thus using Theorem 2.4.1 and (2.2.5) we have

$$\frac{d(V)}{2} \leq \chi^{(\lambda)}(V) \leq d(V) \tag{4.1.4}$$

for $\lambda \geq 1$ with equality on the left if $V$ is a centered set (or $X$ a centered space).

In view of Corollary 4.1.2 it would be helpful for us to find out which spaces are centered. We look first at Euclidean spaces.

If $E$ is an equilateral triangle of side $s$ in $R_2$, then we saw that $\chi^{(\lambda)}(E) = \frac{s}{\sqrt{3}} > \frac{s}{2}$. Thus $R_2$ is not a centered space. In fact we have:

Theorem 4.1.2

Euclidean n-space, $R_n$ for $n > 1$ is not centered.

Proof: Jung [17] shows that if $E$ is a regular simplex of $n+1$ vertices of side $d(E)$ in $R_n$, $n > 1$, its spanning sphere has radius $R = \sqrt{\frac{n}{2(n+1)}}\, d(E)$. Hence Theorem 3.3.4 gives $\chi^{(\lambda)}(E) = \sqrt{\frac{n}{2(n+1)}}\, d(E) > \frac{1}{2}d(E)$.

Thus $R_n$, $n > 1$ is not centered.

Although Euclidean n-space is not centered, $n > 1$, some bounded sets in $R_n$ are centered sets. For these sets, e.g., Case I sets in $R_2$, (see Definition 3.2.2) we have the result $\chi^{(\lambda)}(E) = \frac{d(E)}{2}$, $\lambda \geq 1$.

Theorem 4.1.3

Let $X$ be the real line. Then $X$ is a centered space under any norm.

Proof: Let $V$ be a bounded subset of $X$ with $b = \sup\limits_{x \in V} x$ and $a = \inf\limits_{x \in V} x$. Then $d(V) = \|b - a\|$. Consider $x = \frac{b+a}{2}$. We will show $x$ is a center of $V$. Now $\rho\left(v, \frac{b+a}{2}\right) = \left\|v - \frac{b+a}{2}\right\| = \left\|\frac{2v - (b+a)}{2}\right\|$, or $\rho\left(v, \frac{b+a}{2}\right) \leq \frac{1}{2}(\|v - b\| + \|v - a\|)$. But since for every $v$ there exists a $\mu$ with $0 \leq \mu \leq 1$ such that $v = \mu b + (1 - \mu)a$ we have $\|v - b\| + \|v - a\| = \|b - a\|$. Hence $\rho\left(v, \frac{b+a}{2}\right) \leq \frac{1}{2}\, d(V)$ and $x = \frac{b+a}{2}$ is a center of $V$.

Corollary 4.1.3

Let $V$ be a bounded set on the real line. Then $\chi^{(\lambda)}(V) = \frac{\|b-a\|}{2}$ where $b = \sup_{x \in V} x$, $a = \inf_{x \in V} x$, $\lambda \geq 1$, and $\| \ \|$ denotes the norm defined on the line.

But there exist spaces other than one-dimensional spaces that are centered.

Theorem 4.1.4

The space $R_2^{\infty}$ is centered.

Proof: Let $V$ be a bounded subset of $R_2^{\infty}$. For $v \in V$ we write $v = \left(x^{(1)}, x^{(2)}\right)$. Now

$$d(V) = \sup_{v_i v_j \in V} \|v_i - v_j\| = \sup_{v_i v_j \in V} \max\left(|x_i^{(1)} - x_j^{(1)}|, |x_i^{(2)} - x_j^{(2)}|\right)$$

$$= \max\left(\sup_{v_i, v_j} |x_i^{(1)} - x_j^{(1)}|, \sup_{v_i v_j \in V} |x_i^{(2)} - x_j^{(2)}|\right).$$

Let $Q = \min\left(\sup_{v_i v_j \in V} |x_i^{(1)} - x_j^{(1)}|, \sup_{v_i v_j \in V} |x_i^{(2)} - x_j^{(2)}|\right)$.

Then $V$ is enclosed by a rectangle $G_v$ with sides parallel to the $x^{(1)}$ and $x^{(2)}$ axes and of lengths $d(V)$ and $Q$. The set $G_v$ is a spanning rectangle of $V$; we do not call it a spanning sphere since all points of $G_v$ are not equidistant from its center. This rectangle is unique, for if not, then $V$ would lie in two such rectangles and hence in their intersection, which would be a rectangle with minimum side smaller than $Q$, contradicting the definition of $Q$. Let $O'$ be the center of $G$. Then it is clear that the distance from $O'$ to any point of $V$ is less than or equal to $\frac{d(V)}{2}$. Thus $O'$ is a center for $V$.

It is interesting to note in passing that the intersection of the boundary of $V$ with $G_V$ contains either two points which are at the ends of a diagonal, or at least four points - one point on each side to the rectangle (corner points being counted twice, since they lie simultaneously on two sides of the rectangle).

Corollary 4.1.4

$\chi^{(\lambda)}(V) = \frac{d(V)}{2}$ for $\lambda \geq 1$ for all bounded sets $V$ in $R_2^\infty$.

It is probably true that the proof of Theorem 4.1.4 can be extended to $n$ dimensions and hence Corollary 4.1.4 is valid for all bounded sets in $R_n^\infty$. However we shall not go off on this tangent. Rather we shall stay with the space $R_2^\infty$ and show that the function space $\mathcal{P}_1$ consisting of all real polynomials of degree less than or equal to one defined on an interval, with the uniform norm, is isometric to it.

Let $v_0 = (x_0, y_0) \in R_2^\infty$. To the point $v_0$ we make correspond the polynomial (line) $p_0(x) \in \mathcal{P}_1$, $x \in [a,b]$, joining the points $(a, x_0)$ and $(b, y_0)$. It is clear the correspondence is $1 - 1$. Moreover, in the uniform norm, if $p_0 \in \mathcal{P}_1$ and $p_1 \in \mathcal{P}_1$ we have $\rho(p_0, p_1) = \|p_0 - p_1\| = \max_{x \in [a,b]} |p_0(x) - p_1(x)|$. For linear polynomials this maximum occurs at the endpoints of the interval. Hence $\rho(p_0, p_1) = \max(|x_0 - x_1|, |y_0 - y_1|) = \rho(v_0, v_1)$. Thus the two spaces are isometric and since the Chebyshev constant is a limit of a function involving only the metric of a space we have

$$\chi^{(\lambda)}(V) = \frac{d(V)}{2} \qquad (4.1.5)$$

for $\lambda \geq 1$ for all bounded sets $V$ in $\mathcal{P}_1$ .

This result leads one to speculate whether $\mathcal{P}_1$ is a centered space. Indeed it is, as are spaces of real polynomials defined on an interval for all degrees $n$ in the uniform norm. This is a result derived from the following important theorem found in Kolmogorov-Tihomirov [19].

Theorem 4.1.5

The space $\mathcal{R}(X)$ of real-valued functions defined on an arbitrary set $X$ with the metric $\rho(f,g) = \sup_{x \in X} |f(x) - g(x)|$ is centered.

Proof: Let $V$ be a bounded subset of $\mathcal{R}(X)$ of diameter $d(V)$ . Thus $d(V) = \sup_{f,g \in V} \sup_{x \in X} |f(x) - g(x)|$ . Define $\bar{f}(x) = \sup_{f \in V} f(x)$ and $\underline{f}(x) = \inf_{f \in V} f(x)$ . We see $d(V) = \sup_{x \in X} (\bar{f}(x) - \underline{f}(x))$ . Let $f_0(x) = \frac{1}{2}(\bar{f}(x) + \underline{f}(x))$ . Then $f_0(x)$ is a center of $V$ . For if $f(x) \in V$ , then

$$\underline{f}(x) - f_0(x) \leq f(x) - f_0(x) \leq \bar{f}(x) - f_0(x)$$

or $-\frac{1}{2}(\bar{f}(x) - \underline{f}(x)) \leq (f(x) - f_0(x)) \leq \frac{1}{2}(\bar{f}(x) - \underline{f}(x))$ .

Hence $|f(x) - f_0(x)| \leq \frac{1}{2}(\bar{f}(x) - \underline{f}(x))$ for all $x \in X$ . Thus

$$\rho(f(x),f_0(x)) = \sup_{x \in X} |f(x) - f_0(x)| \leq \frac{1}{2} \sup_{x \in X} (\bar{f}(x) - \underline{f}(x)) = \frac{1}{2} d(V) .$$

This theorem has far-reaching consequences for us because with Corollary 4.1.2 it gives the result:

Theorem 4.1.6

$\chi^{(\lambda)}(V) = \frac{d(V)}{2}$ for $\lambda \geq 1$ where $V$ is any bounded subset of the space $\mathcal{R}(X)$ of real-valued functions defined on an arbitrary set $X$ with sup norm.

The space $\mathcal{C}(X)$ of complex valued functions defined on an arbitrary set $X$ with sup norm is not a centered space. For suppose $V = \left\{f_1(z), f_2(z), f_3(z)\right\}$ such that $f_i(z) = c_i$ , $i = 1,2,3$ for all $z \in X$ , an arbitrary set in $\mathbb{C}$ . Suppose further that $c_i$ are the vertices of an equilateral triangle of side $s$ . Then $V$ has no center, for if $g(z)$ were a center of $V$ , then $\max\limits_{z \in X} |g(z) - f_i(z)| \leq \frac{1}{2} s$ for $i = 1,2,3$ and thus $g(z)$ must be simultaneously inside circles with centers at $c_i$ and radii equal to $\frac{1}{2} s$ . But no point having this property exists. Hence $V$ has no center and $\mathcal{C}(X)$ is not centered.

However, if we restrict our attention to centered subsets of $\mathcal{C}(X)$ we can determine the Chebyshev constant of such sets for certain averages from Theorem 4.1.1. For <u>all</u> averages we can obtain an important upper bound for the Chebyshev constant of centered sets. Since the unit ball of a Banach space is a centered set this will be advantageous.

<u>Theorem 4.1.7</u>

Let $W$ be a centered subset of a metric space $(X, \rho)$ . Then $\chi^A(W) \leq \frac{d(W)}{2}$ for all averages.

<u>Proof:</u> $\tau_n^A(W) = \inf\limits_{\substack{x_i \in X \\ i=1,2,\ldots,n}} \sup\limits_{w \in W} A(\rho(w,x_1), \rho(w,x_2), \ldots, \rho(w,x_n))$ . Let $x_W$ be a center for $W$ . Then $\rho(w,x_W) \leq \frac{d(W)}{2}$ for all $w \in W$ . Thus

$$\tau_n^A(W) \leq \sup_{w \in W} A(\rho(w,x_W), \rho(w,x_W), \ldots, \rho(w,x_W)) = \sup_{w \in W} \rho(w,x_W) \leq \frac{d(W)}{2}$$ and

as $n \to \infty$, $\chi^A(W) \le \frac{d(W)}{2}$.

Corollary 4.1.5

Let U be the unit ball of any normed linear space X. Then $\chi^A(U) \le 1$ for all averages.

We now turn to some work of Hille [16] dealing with the transfinite diameter of the unit ball in certain Banach spaces. In any Banach space X with unit ball U

$$\delta^A(U) \le 2$$

and we have equality in the following spaces:

$C[a,b]$, $L^1(a,b)$, $L^\infty(a,b)$, $\ell$, $m$.

In the $L^p(0,1)$ spaces, Hille obtains a lower bound for the transfinite diameter of the unit ball $U_p$. He shows:

$$\delta^A(U_p) \ge 2\left[\frac{\Gamma(\frac{1}{2}p+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}p+1)}\right]^{\frac{1}{p}} \qquad 1 \le p \le \infty$$

and in the case $1 \le p \le 2$ we have

$$\delta^A(U_p) \ge 2^{\frac{1}{p}}.$$

Hence in all the aforementioned spaces Corollary 4.1.5 implies that we have strict inequality

$$\delta^A(U) > \chi^A(U) \qquad (4.1.6)$$

for the unit ball of the space for all averages.

Another result that holds for all averages is the following theorem furnishing us a lower bound for the Chebyshev constant of the unit ball.

**Theorem 4.1.8**

Let $X$ be a normed linear space of complex (or real) valued bounded functions defined on an arbitrary non-empty set $S$. We suppose further that $X$ contains the constant functions and $\|f\| = \sup_{s\in S} |f(s)|$ for all $f \in X$. Then if $U$ is the unit ball of $X$, $\chi^A(U) \geq \chi^A(Y)$ where $Y$ is the unit disk in the complex plane (or $Y$ is the interval $[-1,1]$ of the real axis).

Proof: Given $\epsilon > 0$, there exists $g_1^*, g_2^*, \ldots, g_n^* \in X$ such that $\tau_n^A(U) + \epsilon > \sup_{f\in U} A(\rho(f,g_1^*), \rho(f,g_2^*), \ldots, \rho(f,g_n^*))$. Take $y \in Y$ arbitrarily. Fix $s_0 \in S$. Since $|y| \leq 1$ and since $X$ contains the constant functions, there exists a function $\hat{f} \in U$ such that $\hat{f}(s_0) = y$. We define a set $\{\eta_i\}$, $i = 1,2,\ldots,n$ by $g_i^*(s_0) = \eta_i$.

Now for every $i$, $i = 1,2,\ldots,n$ we have

$$\rho(\hat{f},g_i^*) = \|\hat{f} - g_i^*\| = \sup_{s\in S} |\hat{f}(s) - g_i^*(s)| \geq |\hat{f}(s_0) - g_i^*(s_0)| = |y - \eta_i| = \rho(y,\eta_i) .$$

Thus

$$\tau_n^A(U) + \epsilon > A(\rho(\hat{f},g_1^*), \rho(\hat{f},g_2^*), \ldots, \rho(\hat{f},g_n^*))$$

and from POSTULATE b) we have

$$\tau_n^A(U) + \epsilon > A(\rho(y,\eta_i), \rho(y,\eta_2), \ldots, \rho(y,\eta_n)) .$$

Since $y$ was arbitrary this last expression holds for all $y \in Y$.

Hence

$$\tau_n^A(U) + \epsilon > \max_{y\in Y} A(\rho(y,\eta_1), (y,\eta_2), \ldots, \rho(y,\eta_n))$$

$$> \min_{\substack{\eta_i\in\mathbb{C} \\ (\text{or } \eta_i\in R_1)}} \max_{y\in Y} A(\rho(y,\eta_1), \rho(y,\eta_2), \ldots, \rho(y,\eta_n)) = \tau_n^A(Y) .$$

Since $\epsilon$ was arbitrary we have, as $n \to \infty$, $\chi^A(U) \geq \chi^A(Y)$ .

Theorem 4.1.8 together with Corollary 4.1.5 and the results of Pólya and Szegö (see appendix) gives

$$\chi^{(\lambda)}(U) = 1 \quad \text{for} \quad \lambda \geq 0 \qquad (4.1.7)$$

for any normed linear space X of complex-valued bounded functions satisfying the hypothesis of Theorem 4.1.8.

In the real case one can infer $\chi^{(\lambda)}(U) = 1$ for $\lambda \geq 1$ , a result we have already obtained in Corollary 4.1.1.

Section 4.2 is devoted to a concept discussed in Singer [28]. Although it appears in a different context, Chebyshev centers, as the name might suggest, is also related to the ideas developed in the preceding chapters as well as to the concept of centered sets.

## 4.2. Chebyshev Centers

Let X be a metric space and V a bounded subset of X . We examine the value $\tau_n^A(V)$ for $n = 1$ . Using the definition of the footnote for (2.3.1) (p.37), we have $\tau_1^A(V) = \inf_{x_1 \in X} \sup_{x \in V} A(\rho(x,x_1))$ or $\tau_1^A(V) = \inf_{x_1 \in X} \sup_{x \in V} \rho(x,x_1)$ . Since $\tau_1^A(V)$ is independent of the average we shall write it as $\tau_1(V)$ .

### Definition 4.2.1

A point $x_1^* \in X$ such that $\tau_1(V) = \sup_{x \in V} \rho(x,x_1^*)$ is called a Chebyshev center for V . The value $\tau_1(V)$ is called a Chebyshev radius for V .

The terminology of Definition 4.2.1 is natural since we see that

$x_1^*$ is the center of a " sphere" of smallest radius $\tau_1(V)$ which encloses $E$ . Indeed $x_1^*$ is a center of a spanning sphere of Definition 3.3.1 and $\tau_1(V)$ is its radius. Such a point $x_1^*$ defined by the above definition is of course a Chebyshev point of order one for $V$ , which is the same for all average functions.

The relationship between centers of sets and Chebyshev centers is as follows:

Theorem 4.2.1

Let $W$ be a centered subset of a metric space $(X,\rho)$ . Then if $x_W$ is a center for $W$ , $x_W$ is a Chebyshev center for $W$ .

Proof: Suppose $x_W$ is not a Chebyshev center for $W$ . Then $\sup_{x \in W} \rho(x,x_W) > \tau_1(W)$ . Since $W$ is centered we have from Theorem 4.1.1, $\tau_1(W) = \frac{d(W)}{2}$ . Hence $\sup_{x \in W} \rho(x,x_W) > \frac{d(W)}{2}$ . Thus $x_W$ is not a center for $W$ .

Theorem 4.2.2

Any Chebyshev center of a centered set $W$ in a metric space is a center.

Proof: Since $\sup_{x \in W} \rho(x,x_1^*) = \tau_1(W)$ and $\tau_1(W) = \frac{d(W)}{2}$ from Theorem 4.1.1, we have $\rho(x,x_1^*) \leq \frac{d(W)}{2}$ for all $x \in W$ and $x_1^*$ is a center.

We note that the hypothesis of Theorem 4.2.2 cannot be weakened to sets $V$ which are merely bounded. For example, if $V$ were an equilateral triangle, then its circumcenter is a Chebyshev center, but not a center!

When the set in question is compact the existence of $x_1^*$ is assured. But if $V$ is merely bounded, one has to look at the existence question. We have the following theorem of A.L. Garkavi [8]:

Theorem 4.2.3

Let X be a Banach space with the property that there exists a projection $p : X^{**} \to X$ of norm one. Then for every bounded set $V \subset X$ there exists a Chebyshev center for V .

The existence of such a projection as described in the hypothesis is known for Banach spaces which are equivalent to conjugate spaces. This does not characterize conjugate spaces however, for such projections exist in spaces which are not the conjugate space of any Banach space. See Ruston [26].

Of more interest to us is the uniqueness question of $x_1^*$ . In n-dimensional Euclidean space we showed that the spanning sphere is unique. This is not true in general. We have the following uniqueness theorem of A.L. Garkavi [8].

Theorem 4.2.4

In order that every bounded set V in a normed linear space X has at most one Chebyshev center, it is necessary and sufficient that the space X be uniformly convex [1] in every direction.

Theorems 4.2.3 and 4.2.4 imply the existence and uniqueness of a Chebyshev center for any bounded set in a Hilbert space.

In Euclidean spaces the Chebyshev center played an important role since the Chebyshev points of order n for E , for all n , coincided with the Chebyshev center for λ-th power averages, $\lambda \geq 1$ .

A space which is not uniformly convex is $R_2^\infty$ . Thus there exist

---

[1] A normed linear space X is uniformly convex in every direction if for every $\epsilon > 0$ and for every $x \in X$ , there exists a $\delta(\epsilon,x) > 0$ such that if $\|y\| = \|z\| = 1$ and $y - z = \lambda x$ and $\|y+z\| > 2-\delta$ , then $|\lambda| \leq \epsilon$ .

bounded sets $V$ in $R_2^\infty$ which have more than one Chebyshev center. Consider

Example 4.2.1

Let $X = R_2^\infty$ and $V = [-1,1]$ on $x^{(1)}$ axis. Now

$$\tau_1(V) = \min_{x_1 \in X} \max_{x \in V} \|x - x_1\| = \min_{x_1 \in X} \max_{x \in V} \max\left(|x^{(1)} - x_1^{(1)}|, |0 - x_1^{(2)}|\right) .$$

For $x_1$ not on $x^{(2)}$ axis, the maximum over all $x \in V$ of the above expression is strictly greater than $1$ . For $x_1$ any point on the $x^{(2)}$ axis in the segment from $-1$ to $1$ , we have the above expression equal to $1$ and thus $\tau_1(V) = 1$ . Therefore all points in the segment from $-1$ to $1$ on the $x^{(2)}$ axis are Chebyshev centers for $V$ . We observe that in this case not only is $x_1^*$ not unique, but it does not necessarily lie in $K(V)$ , the closed convex hull of $V$ . The spanning sphere $S_V$ of the set $V$ consists of any square with sides parallel to the $x^{(1)}$ and $x^{(2)}$ axes of length two and with center on the segment from $-1$ to $1$ on the $x^{(2)}$ axis. The spanning rectangle $G_V$ of $V$ , as discussed in the proof of Theorem 4.1.4, is a degenerate one, consisting merely of $V$ itself with center at the origin.

We close this section with another theorem of Garkavi [9].

Theorem 4.2.5

In order that every bounded set $V$ in a Banach space $X$ have a Chebyshev center $x_1^*$ with $x_1^* \in K(V)$ , it is necessary and sufficient that $X$ be a Hilbert space or $\dim X = 2$ .

We shall see how this modified "convex hull theorem," pertaining to only the Chebyshev point of order one, relates to the ideas that will be discussed in Section 4.3. Again in this section we use concepts discussed in another context by Singer [28].

## 4.3 Closest Points and Fejér

Definition 4.3.1

Let $X$ be a normed linear space. Let $x,y$ be two elements of $X$ and let $Z$ be any subset of $X$. We say $y$ is point-wise closer to $Z$ if $\|y - z\| < \|x - z\|$ for each $z \in Z$. If there exists no $y$ such that $\|y - z\| < \|x - z\|$ for each $z \in Z$, then we say $x$ is a closest point to $Z$; i.e., a closest point is one for which there is no point-wise closer point.

We shall denote by $C(Z)$ the set of closest points of $Z$. From the definition we see that if $x \in Z$, then $x \in C(Z)$. Thus we have

$$Z \subset C(Z) \tag{4.3.1}$$

The following theorem was proved by Fejér [5] in 1922. It should not be confused with Fejér's Principle, which we shall comment on presently.

Theorem 4.3.1

Let $E$ be a compact set in the Euclidean plane. Then $C(E) = K(E)$ where $K(E)$ is the closed convex hull of $E$.

Fejér's restriction to compact sets is unnecessary. Theorem 4.3.1 has been generalized in the form:

Theorem 4.3.2

If $Z$ is any subset of a Hilbert space, then $C(Z) = K(Z)$.

The proof of this theorem is quoted in Phelps [24].

Definition 4.3.2

Let $X$ be a normed linear space. A subset $Z$ of $X$ will be called a strong Fejér set if $C(Z) = K(Z)$. If all subsets of $X$ are strong Fejér sets, then we shall say that the space $X$ is <u>strongly Fejér</u>.

Theorem 4.3.2 shows that Hilbert space is, in our terminology, strongly Fejér.

Now let us look at Fejér's Principle. The Principle states that for a certain subset $Z$ of $X$, and for every $x \in X \backslash Z$, there exists a $y \in X$ such that $\|y - z\| < \|x - z\|$ for each $z \in Z$. (In Lemma 1.2.1, $X$ was $\mathbb{C}$ and $Z$ was a compact convex subset of $X$.) But in view of Definition 4.2.1, Fejér's Principle states that if $x \notin Z$ then $x \notin C(Z)$, or

$$C(Z) \subset Z \tag{4.3.2}$$

In view of (4.3.1) this is equivalent to $C(Z) = Z$.

<u>Definition 4.3.3</u>

A subset $Z$ of a normed linear space $X$ will be called a <u>principle Fejér set</u> if $C(Z) = Z$.

It is obvious that not all subsets of a space can be principle Fejér sets. Certainly all open proper subsets are not principle Fejér sets.

From the definitions one can easily show the following:

<u>Proposition 4.3.1</u>

If a normed linear space $X$ is strongly Fejér, then every closed convex subset of $X$ is a principle Fejér set.

The hypothesis of Proposition 4.3.1 is stronger than need be. Consider the following definition:

<u>Definition 4.3.4</u>

A subset $Z$ of a normed linear space $X$ will be called a <u>weak</u>

<u>Fejér set</u> if $C(Z) \subset K(Z)$ . If all subsets of X are weak Fejér sets, then we shall say that the space X is <u>weakly Fejér</u>.

From the definitions, each principle Fejér set is a weak Fejér set. It is easy to prove

<u>Proposition 4.3.2</u>

If a normed linear space X is weakly Fejér, then every closed convex subset of X is a principle Fejér set.

A space X , which according to our definition is weakly Fejér, is said by Phelps [24] to possess Property F. He proves the three interesting theorems (written in our terminology).

<u>Theorem 4.3.3</u>

Let X be a normed linear space of dimension $\geq 3$ which is weakly Fejér. Then X is a Hilbert space.

<u>Theorem 4.3.4</u>

Let X be a two-dimensional normed linear space. Then X is strictly convex [1] if and only if X is weakly Fejér.

<u>Theorem 4.3.5</u>

If V is a bounded subset of a strictly convex two dimensional space, then V is a strong Fejér set.

Fejér's Principle implies the convex hull theorem in the following manner: Let E be a compact subset of a normed linear space X , with K(E) its closed convex hull. If K(E) is a principle Fejér set, then the Chebyshev points of E of order n , for all n , and for all averages are located in K(E) .

---

[1] A normed linear space X is said to be strictly convex if for every x and y in X such that $\|x\| = \|y\| = 1$ and $x \neq y$ we have $\left\| \frac{x+y}{2} \right\| < 1$ .

From Proposition 4.3.2 we thus have a convex hull theorem for all spaces which are weakly Fejér. In particular Theorem 4.3.2 with Proposition 4.3.1 implies the convex hull theorem for all compact subsets of a Hilbert space, a result we have proved in Theorem 2.3.4. In addition we have a new result derived from Theorem 4.3.4 and Proposition 4.3.2. Together they imply the convex hull theorem for all compact subsets of a two-dimensional strictly convex normed linear space.

We now examine how the theorems of Phelps relate to the modified convex hull theorem of Garkavi, Theorem 4.2.5. Since the convex hull theorem is satisfied in weakly Fejér spaces, surely in such spaces we have $x_1^* \in K(E)$ .

Theorem 4.3.2 and 4.3.4 provide sufficient conditions for a space to be weakly Fejér. Theorems 4.3.3 and 4.3.4 provide necessary conditions for a space to be weakly Fejér.

The sufficiency conditions in Theorems 4.3.2 and 4.3.4 (that a space be either Hilbert or two-dimensional strictly convex) are sufficiency conditions for Theorem 4.2.5. They are not necessary since the strict convexity in two-dimensional spaces insures the convex hull theorem for Chebyshev points of order $n$ , for every $n$ , but is evidently not necessary for just Chebyshev points of order one.

The necessary conditions in Garkavi's Theorem 4.2.5 (that a space be either Hilbert or two-dimensional) are necessary conditions for a space to be weakly Fejér. They are not sufficient since the strict convexity in two-dimensional spaces is lacking.

| Set | Transfinite Diameter | | Chebyshev Constant | |
|---|---|---|---|---|
| $\{-1,1\}$ | $2^{1-\frac{1}{\lambda}}$ | $\lambda > 0$ | $1$ | $\lambda \geqq 1$ |
| | | | $2^{1-\frac{1}{\lambda}}$ | $0 < \lambda \leqq 1$ |
| | $0$ | $\lambda \leqq 0$ | $0$ | $\lambda \leqq 0$ |
| $[-1,1]$ | $2^{1-\frac{1}{\lambda}}$ | $\lambda \geqq 1$ | $1$ | $\lambda \geqq 1$ |
| | $\left\{\frac{\Gamma\left(\frac{1+\lambda}{2}\right)\Gamma\left(1-\frac{\lambda}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\right\}^{\frac{1}{\lambda}}$ | $0 \leqq \lambda \leqq 1$ | $\left\{\frac{\Gamma\left(\frac{1+\lambda}{2}\right)\Gamma\left(1-\frac{\lambda}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\right\}^{\frac{1}{\lambda}}$ | $-1 \leqq \lambda \leqq 1$ |
| | $0$ | $\lambda \leqq -1$ | $0$ | $\lambda \leqq -1$ |
| $\{z \in \mathbb{C} \mid \lvert z\rvert = 1\}$ | $2^{1-\frac{1}{\lambda}}$ | $\lambda \geqq 2$ | $1$ | $\lambda \geqq 0$ |
| | $\frac{\Gamma(1+\lambda)^{\frac{1}{\lambda}}}{\Gamma\left(1+\frac{\lambda}{2}\right)^{\frac{2}{\lambda}}}$ | $-1 \leqq \lambda \leqq 2$ | $\frac{\Gamma(1+\lambda)^{\frac{1}{\lambda}}}{\Gamma\left(1+\frac{\lambda}{2}\right)^{\frac{2}{\lambda}}}$ | $-1 \leqq \lambda \leqq 0$ |
| | $0$ | $\lambda \leqq -1$ | $0$ | $\lambda \leqq -1$ |
| $\{z \in \mathbb{C} \mid \lvert z\rvert \leq 1\}$ | $2^{1-\frac{1}{\lambda}}$ | $\lambda \geqq 2$ | $1$ | $\lambda \geqq 0$ |
| | $\frac{\Gamma(1+\lambda)^{\frac{1}{\lambda}}}{\Gamma\left(1+\frac{\lambda}{2}\right)^{\frac{2}{\lambda}}}$ | $0 \leqq \lambda \leqq 2$ | $\left(\frac{\frac{\pi\lambda}{2}}{\sin\frac{\pi\lambda}{2}}\right)^{\frac{1}{\lambda}}$ | $-2 \leqq \lambda \leqq 0$ |
| | $0$ | $\lambda \leqq -2$ | $0$ | $\lambda \leqq -2$ |

| Set | Transfinite Diameter | | Chebyshev Constant | |
|---|---|---|---|---|
| $\{(x_1,x_2,x_3) \in R_3 \mid x_1^2+x_2^2+x_3^2 = 1\}$ | $2^{1-\frac{1}{\lambda}}$ | $\lambda \geqq 2$ | $1$ | $\lambda \geqq -1$ |
| | $2\left(1+\frac{\lambda}{2}\right)^{-\frac{1}{\lambda}}$ | $-2 \leqq \lambda \leqq 2$ | $2\left(1+\frac{\lambda}{2}\right)^{-\frac{1}{\lambda}}$ | $-2 \leqq \lambda \leqq -1$ |
| | $0$ | $\lambda \leqq -2$ | $0$ | $\lambda \leqq -2$ |
| $\{(x_1,x_2,x_3) \in R_3 \mid x_1^2+x_2^2+x_3^2 \leq 1\}$ | $2^{1-\frac{1}{\lambda}}$ | $\lambda \geqq 2$ | $1$ | $\lambda \geqq -1$ |
| | $2\left(1+\frac{\lambda}{2}\right)^{-\frac{1}{\lambda}}$ | $-1 \leqq \lambda \leqq 2$ | $\left\{\frac{\Gamma\left(\frac{3+\lambda}{2}\right)\Gamma\left(1-\frac{\lambda}{2}\right)}{\Gamma\left(\frac{3}{2}\right)}\right\}^{\frac{1}{\lambda}}$ | $-3 \leqq \lambda \leqq -1$ |
| | $0$ | $\lambda \leqq -3$ | $0$ | $\lambda \leqq -3$ |

## BIBLIOGRAPHY

1 Ankeny, N.C. and T.J. Rivlin, On a theorem of S. Bernstein, Pac. Journ. of Math., 5 (1955) pp. 849-852.

2 Cheney, W. and A.A. Goldstein, Proximity maps for convex sets, Proc. Amer. Math. Soc. 10 (1959) pp. 448-450.

3 Courant, R. and H. Robbins, What is Mathematics? Oxford University Press, London, New York (1953).

4 Davis, P.J., Interpolation and Approximation, Blaisdell, New York (1963).

5 Fejér, L., Über die Lage der Nullstellen von Polynomen die aus Minimumforderungen gewisser Art entspringen, Math. Ann. 85 (1922) pp. 41-48.

6 Fekete, M., Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungun mit ganzzahligen Koeffizienten, Math. Z. 17 (1923) pp. 228-249.

7 Frostman, O., Potentiel d'équilibre et capacité des ensembles avec quelques applications à la théorie des fonctions, Thèse, Lund (1935) 118 pp.

8 Garkavi, A.L., On the optimal net and best cross-section of a set in a normed space, (Russian), Isv. Akad. Nauk SSSR Ser. Mat. 26 (1962) pp. 87-106.

9 Garkavi, A.L., On the Chebyshev center and convex hull of a set, (Russian), Uspehi Mat. Nauk 19 (1964), No. 6 (120) pp. 139-145.

10 Goluzin, G.M., Geometrische Funktionentheorie, Deutscher Verlag der Wissenschaften, Berlin (1957); Translated: Geometric Theory of Functions of a Complex Variable, A.M.S., Providence, R.I. (1969).

11 Hardy, G.H., J.E. Littlewood, and G. Pólya, Inequalities, Cambridge Univ. Press, Cambridge, England (1934).

12 Hayman, W.K., Transfinite Diameter and its Application, Institute of Mathematical Sciences, Madras (1965).

13 Hille, E., Remarks on transfinite diameters, Gen. Topology Rel. Mod. Analysis and Algebra, Proc. Symp. Prague, (1961) pp. 211-220; Czechoslovak Acad. Sci., Prague (1962).

14. Hille, E., Analytic Function Theory, II, Ginn, Boston (1962).

15 Hille, E., A note on transfinite diameters, J. Analyse Math., 14 (1965) pp. 209-223.

16 Hille, E., Topics in Classical Analysis, Lectures on Modern Math., Vol. 3, Wiley & Sons, New York (1965) pp. 1-57.

17 Jung, H.W.E., Über die kleinste Kugel, die eine räumliche Figur einschliest, J. fur Math. 123 (1901) pp. 241-257.

18 Kolmogorov, A.N., Sur la notion de la moyenne, Rend. Acad. Lincei (6) 12 (1930) pp. 388-391.

19 Kolmogorov, A.N. and U.M. Tihomirov, ϵ-entropy and ϵ-capacity of sets in function spaces, (Russian), Uspehi Mat. (N.S.) 14 (1959), no. 2 (86) pp. 3-86; Translated: American Math. Soc. Translations, Series 2, vol. 17 (1961) pp. 277-364.

20 Leja, F., Sur la définition du diamètre et de l'ecart transfini d'un ensemble, Ann. Soc. Polon. Math., 12 (1934) pp. 29-34.

21 Leja, F., Sur une suite de polynomes et la représentation conforme d'un domain plan quelconque sur le cercle, Ann. Soc. Polon. Math., 14 (1936) pp. 116-134.

22 Leja, F., Généralisation de certaines fonctions d'ensemble. Ann. Soc. Polon. Math., 16 (1937) pp. 41-52.

23 Nagumo, M., Über eine Klasse der Mittlewerte, Jap. J. Math., 7 (1930) pp. 71-79.

24 Phelps, R.R., Convex sets and nearest points, II, Proc. Amer. Math. Soc. 9 (1958) pp. 867-873.

25 Pólya, G. and G. Szegö, Über den transfiniten Durchmesser (Kapazitatskonstante) von ebenen und räumlichen Punktmengen, J. Reine Angew Math. 165 (1931) pp. 4-49.

26 Ruston, A.F., A note on the greatest crossnorm, Proc. Amer. Math. Soc. 13 (1962) pp. 828-829.

27 Shisha, O., A remark on Fejér's theorem on the convex hull of a point set, (Hebrew), Riveon Lematematika, (1955) pp. 75-77.

28 Singer, I., Best Approximation in Normed Vector Spaces by Elements of Vector Subspaces, Editura Academiei Republicii Socialiste Romania, Bucharest (1967).

29 Steinhaus, H., Mathematical Snapshots, Oxford University Press, New York (1960).

30 Szegö, G., Bemerkung zu einer Arbein von Herrn M. Fekete: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, Math. Z., 21 (1924) pp. 206-208.

31 Szegö, G., Orthogonal Polynomials, (Revised Edition), A.M.S. New York (1959).

32 Tsuji, M., Potential Theory in Modern Function Theory, Maruzen, Tolyo (1959).

33 Yaglom, I., Convex Figures, Holt, Rinehart, and Winston, New York (1961).

# AUTOBIOGRAPHY

Susan Loeserman Friedman was born on February 26, 1934 in Minneapolis, Minnesota, but grew up in New York City. Here she attended the High School of Music and Art and Queens College, where she was awarded a Bachelor of Science degree, magna cum laude in 1955. During a six year stay in the environs of Washington, D.C., she earned a Master of Arts degree in 1957 from the University of Maryland. On her return to New York, Mrs. Friedman taught at both Queens College and Hunter College, and enrolled in the doctoral program at The City University of New York in the first class in 1964.

Since 1952 she has been married to Stanley David Friedman, a writer-producer of television documentaries, and they are the parents of Deborah Caryn, Daniel Seth, and Michael Frederic Friedman.