71-19,768

TANNER, Robert Michael, 1946-
CONTRIBUTIONS TO THE SIMPLEX CODE CONJECTURE.

Stanford University, Ph.D., 1971
Mathematics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

CONTRIBUTIONS TO THE SIMPLEX CODE CONJECTURE

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FUFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

By

Robert Michael Tanner

December 1970

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Thomas M Cover

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert M. Gray

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Richard A. Pantell

Approved for the University Committee
on the Graduate Division:

Lincoln E Moses

Dean of the Graduate Division

ABSTRACT

In this thesis we examine the signal design problem for the infinite bandwidth Gaussian channel. Attention is focused on the well-known simplex code conjecture which posits the optimality of the signal set whose vectors are the vertices of a regular n-dimensional simplex.

Using geometrically inspired methods, we present a simplified proof of the inequality developed by Landau and Slepian in their 1966 paper. The key elements permitting the simplification are a spherical projection and a lemma which shows that the requirements imposed on the optimum solution can be relaxed. In the course of the exposition the difficulties preventing application of the inequality in spaces of dimension greater than three are clarified.

Next, a series of conjectures on the properties of the regular simplex are proposed. To compare the regular simplex with any other region formed by a collection of n hyperplanes in n-space, we define a function $t(r)$ which is the $(n-1)$-dimensional measure of the intersection of the shell $|X| = r$ with each configuration. The first conjecture asserts that the function corresponding to the regular simplex and the function corresponding to the arbitrary region have only one crossing. The second asserts that if the integrals from 0 to r_o of the two are equal, the simplex function is less than or equal to the other for all r larger than r_o. The third assumes that all of the hyperplanes of the competing region and those forming the regular simplex are the same distance from the origin and postulates that the simplex function is in this case less than or equal to the other for all r. We then prove that a sufficient condition for the latter is that, given $n-1$ points and a shell $|X| = r$

in $(n-1)$-dimensional space, there is at most a unique position for an n-th point such that the intersection of the shell with the polytope defined by the n points has its center of mass at the origin.

Returning to the conjecture itself, we formulate and prove a "linearized" version of the problem. The linearized conjecture is that if the tips of the constrained code vectors are regarded as electrons which repel one another according to a specified force law, the regular simplex is the minimum energy configuration. Finally, we comment on the geometrical interpretation of the relation between optimality at low signal-to-noise ratios λ and the mean width of the polytope with the signal set as vertices. In addition, a search through the literature indicates that a much-cited proof that the regular simplex maximizes the mean width may not exist, and thus optimality as λ tends to zero appears to remain an open question.

CONTENTS

ACKNOWLEDGEMENT

I wish to thank Professors Thomas M. Cover, Robert M. Gray, and Richard A. Pantell for their constructive comments on this manuscript. To Professor Cover I am especially grateful for his continued moral support in times of great need.

To the National Science Foundation I am deeply indebted for the freedom of inquiry which their generous and sustained financial assistance has afforded me.

Chapter I

INTRODUCTION

A. The Problem and its History

This thesis is concerned with a problem of communication theory sometimes called the simplex code conjecture. The well-known model from which it arises is that of a channel with additive noise across which one of M equally likely messages is to be transmitted. Assuming the criterion for judging the quality of transmission is the probability of error of the system, we wish to determine what signal set, S_i, $i=1,\ldots,M$, will permit the receiver to best distinguish one signal from another and thus enable him to decide which message was sent with the minimum probability of error.

The first step in the reduction of the problem to mathematically tractable form was made when Shannon [1] and Kotel'nikov [2] introduced the geometric representation of signals. In the geometric context, each of the signals can be viewed as a point in Euclidean n-space, where the dimension n is related to the bandwidth of the channel. The projections of the signal vector along the axes are the coefficients of an expansion of the signal in terms of an orthonormal basis for the signal space, and consequently the length of the signal vector is proportional to its energy. When S_i is sent, the received vector is $\underline{\overline{X}} = S_i + \underline{\overline{Z}}$, the original signal plus a random noise vector distributed according to some probability density function $p_Z(Z) = \Pr\{\underline{\overline{Z}} = Z\}$.[1] To minimize the probability of

1. For convenience, plain capitals will be used to denote both vectors and functions. When deemed necessary for clarity, an arrow $(\vec{X})$ will be added to vectors. The double bar $\underline{\overline{\cdot}}$ signifies random variable. Also, we shall use the common shorthand for probabilities $\Pr\{Z\} \equiv \Pr\{\underline{\overline{Z}} = Z\}$.

error, the optimum receiver (see, for example, Wozencraft and Jacobs [3]) will make the Bayes decision, deciding message i was sent if

$$\Pr\{S_i|X\} = \frac{p_z(X-S_i)\Pr\{S_i\}}{\Pr\{X\}} \geq \frac{p_z(X-S_j)\Pr\{S_j\}}{\Pr\{X\}} \quad \text{for all} \quad j. \tag{I.1}$$

Denoting by $\mathfrak{R}_i$ the set of X's for which the above inequality holds,[1] the probability of error can then be written

$$P_e = 1 - \sum_i \int_{\mathfrak{R}_i} p_z(X-S_i) \Pr\{S_i\} dV \tag{I.2}$$

where dV is an n-dimensional volume element.

As long as the signals are distinct, it will always be possible to achieve an arbitrarily small probability of error merely by increasing the transmitter power to scale up the signals uniformly. Realistically, however, there is almost inevitably a positive cost associated with energy, and thus we are led to impose some energy constraint. Several such constraints and their relations are discussed by Shannon [4]. The applicable constraint for any particular problem depends, of course, on the physical transmitter design. For mathematical simplicity we will adopt the equal energy restriction; specifically, we will assume that $|\hat{S}_i| = 1$ for all i, where $|\cdot|$ denotes the usual Euclidean norm.[2]

An additional simplifying assumption frequently made is that the noise function is spherically symmetrical and monotonically decreasing. More precisely,

$$p_z(Z) = f(|Z|) \tag{I.3}$$

1. Script letters shall be reserved for sets or regions.
2. In accordance with the notation of Landau and Slepian [5], we shall henceforth use the caret (^) to denote unit vector.

where $f(\cdot)$ is a monotonically decreasing function of its argument. White Gaussian noise can be seen to be a member of this class. Again, for mathematical expediency, we will limit our attention to noise functions of this form.

Finally, we wish to examine the situation where there is no bandwidth constraint, or, geometrically, where the dimension of the space, n, can be taken to be arbitrarily large. Since M points are always contained in an $(M-1)$-dimensional subspace, n need be no bigger than $M-1$, because if this subspace does not contain the origin, then the M points lie on an $(M-1)$-dimensional hypersphere of radius less than 1. Increasing the length of each signal can only decrease the probability of error, so we may assume that the subspace does contain the origin.

The problem can now be summarized as follows: Given M unit vectors $\hat{S}_i$, $i=1,\ldots,M$, in $(M-1)$-space and a probability density function $p_Z(Z)$ which is monotonically decreasing in $|Z|$, what set of $\hat{S}_i$ minimizes

$$P_e = 1 - 1/M \int_{\mathcal{R}_1} \max_i p_Z(X-\hat{S}_i)\,dV \tag{I.4}$$

where (I.1) and the equiprobable signal assumption have been used to simplify (I.2). After a moment's contemplation it is easy to see that the most likely candidate is the set with the greatest degree of symmetry, namely, the set whose vertices form an equilateral triangle in 2-space, a regular tetrahedron in 3-space, or, in general, a regular simplex [6]. Denoting by $\underline{S}$ the matrix with $\hat{S}_1,\ldots,\hat{S}_M$ as columns, the regular simplex can be defined by

$$\underline{\alpha} = \underline{S}^T\underline{S} = \{\alpha_{ij}\} \qquad \alpha_{ij} = 1 \qquad j=i$$
$$\alpha_{ij} = \frac{-1}{M-1} \qquad j \neq i\ . \tag{I.5}$$

The optimality of this set, which is the assertion of the "simplex code conjecture," has gone unproved for the past twenty years.

As far as we know, the problem was first suggested by Kotel'nikov, when, at the end of a section entitled "The optimum system for signals with many discrete values," he concludes that the regular simplex signals " . . . are the optimum systems (at least among the family of systems of equidistant signals)." Gilbert [7] made numerical evaluations of the probability of error for the simplex codes and compared them with other signal sets. Asymptotic bounds for the optimal signal sets in Gaussian noise were obtained by Shannon [4] using geometrically inspired inequalities. Balakrishnan [8] developed a new expression for the probability of correct decoding which enabled him to show that for Gaussian noise the regular simplex is a local maximum (local minimum for probability of error) at all signal-to-noise ratios λ as well as being the only possible maximum independent of λ in some interval. In a later work [9] he demonstrated that as λ tends toward infinity, the optimum signal set must converge to the regular simplex, and that optimality as λ tends toward zero is equivalent to proving that the regular simplex maximizes the mean width of all polytopes with M vertices lying on the $(M-1)$-dimensional unit hypersphere (see Chapter V of this thesis). Extensive additional efforts along the line of attack pursued by Balakrishnan may be found in the recent book by Weber [10].

In 1966, Wyner [11] refined the bounds indicated by Shannon [4] to prove that the error-exponent for the simplex codes is optimal. Shortly after Wyner's paper was written, Landau and Slepian [5] announced a proof of the conjecture for general spherically symmetric monotonically decreasing noise functions and arbitrary number of signals M. The proof was based on the generalization of a theorem of Fejes-Toth [12] which deals with the geometry of the optimal decoding regions. Their success was ephemeral, however. In 1968 Farber [13] revealed that the theorem described the entire signal geometry only in spaces of dimension $n \leq 3$. He also provided an alternative proof of the Fejes-Toth theorem for $n = 3$ by mapping a proof given by Schaffner and Kreiger [14] for the incoherent two-dimensional channel into the coherent three-space. This mapping is further discussed in a recent paper by Blachman [15].

In this thesis we take up the quest for a proof of global optimality using geometric methods most closely related to those of Landau and Slepian. In Chapter II we reprove the extension of Fejes-Toth's theorem employing techniques which simplify substantially the mathematics involved. In so doing, the difficulties preventing application of the inequality in spaces of dimension greater than three are clarified. Chapter III presents a hierarchy of conjectures on the properties of the regular simplex which seem to be implicit in many of the approaches which the problem suggests. The conjectures are ordered and compared in an effort to help future researchers recognize more rapidly whether an apparently new path will be fruitful. Some new sufficient conditions are also given. In Chapter IV we formulate and prove a linearized version of the simplex code conjecture. The linearized problem is seen to be one of

" n-dimensional electrostatics," where the tips of the constrained code vectors take the role of electrons which repel one another according to a specified force law. The essential difference between the versions is that the force between two vectors is unaffected by the presence of a third, whereas the associated probability of error is changed. Finally, Chapter V comments on the geometric interpretation of the relation between optimality at low signal-to-noise ratios λ and the mean width of the polytope with the signal set as vertices. A search through the literature indicates that a much cited proof that the regular simplex maximizes the mean width may not exist, and thus optimality as λ tends to zero appears to remain an open question.

B. Background Approaches

There are numerous ways of subdividing the simplex conjecture in an effort to simplify analysis. To provide a setting for our own work, we briefly discuss some of the reductions proposed by previous authors.

Using the equally likely message assumption and the form of the noise function, we can rewrite the decision region $\mathfrak{R}_i$ as

$$\mathfrak{R}_i = \{X: |X - \hat{S}_i| < |X - \hat{S}_j| \quad \text{for all} \quad j \neq i\} \tag{I.6}$$

Note that the boundaries between regions have zero measure and therefore need not be assigned. Since the $\hat{S}_i$ are all unit vectors, $\mathfrak{R}_i$ can also be expressed

$$\mathfrak{R}_i = \{X: X \cdot \hat{S}_i > X \cdot \hat{S}_j \quad \text{for all} \quad j \neq i\} \tag{I.7}$$

It is easy to verify that $\mathfrak{R}_i$ is convex and radially invariant. Equation (I.2) then becomes

$$P_e = 1 - \frac{1}{M} \sum_i \int_{\mathcal{R}_i} p_z(X - \hat{S}_i)dV \tag{I.8}$$

The integration can be performed by integrating first over the surface of the hypersphere of radius r and then over r:

$$P_e = 1 - \int_0^\infty \frac{1}{M} \sum_i \int_{\mathcal{R}_i \cap \mathcal{S}_o(r)} p_z(X - \hat{S}_i)dSdr \tag{I.9}$$

where we define $\mathcal{S}_o(r)$ to be $\{X: |X| = r\}$. Landau and Slepian define the function $U(r)$ to be

$$U(r) = \frac{1}{M} \sum_i \int_{\mathcal{R}_i \cap \mathcal{S}_o(r)} p_z(X - \hat{S}_i)dS \tag{I.10}$$

They then wish to show that for each r $U(r)$ is maximized if the signal set is a regular simplex. It is known, by what amounts to the familiar Neyman-Pearson Lemma of statistics, that, were it possible, the best decision regions for $r = 1$ would be a collection of equal size caps

$$\mathcal{C}_i(\varphi) \equiv \{X: X \cdot S_i > \cos \varphi\} \tag{I.11}$$

where the angle φ of the caps is chosen to satisfy

$$M \int_{\mathcal{C}_i} dS = \int dS = \text{Area of unit hypersphere.} \tag{I.12}$$

(This is similar to Shannon's sphere-packing bound as in, e.g., [4].) Because the problem is of the same form for every r, we can assume without loss of generality that $r = 1$. Since such a collection of caps cannot be disjoint, Landau and Slepian, in the manner of Fejes-Toth, start with a set of caps with a larger angle φ and attempt to show that, in effect, the regular simplex decoding regions $\mathfrak{R}_i$ represent the best modification of these caps.

Another direction suggested by Farber is to observe that any noise function of the required form can be uniformly approximated by a series of positive ball densities

$$p_z(Z) = f(|Z|) = k(r_o) \qquad |Z| < r_o$$
$$= 0 \qquad |Z| \geq r_o \quad . \tag{I.13}$$

Thus it suffices to show optimality of the regular simplex for this class of functions. We shall refer to this problem as the ball density conjecture. As before, it is also possible to consider the optimization at

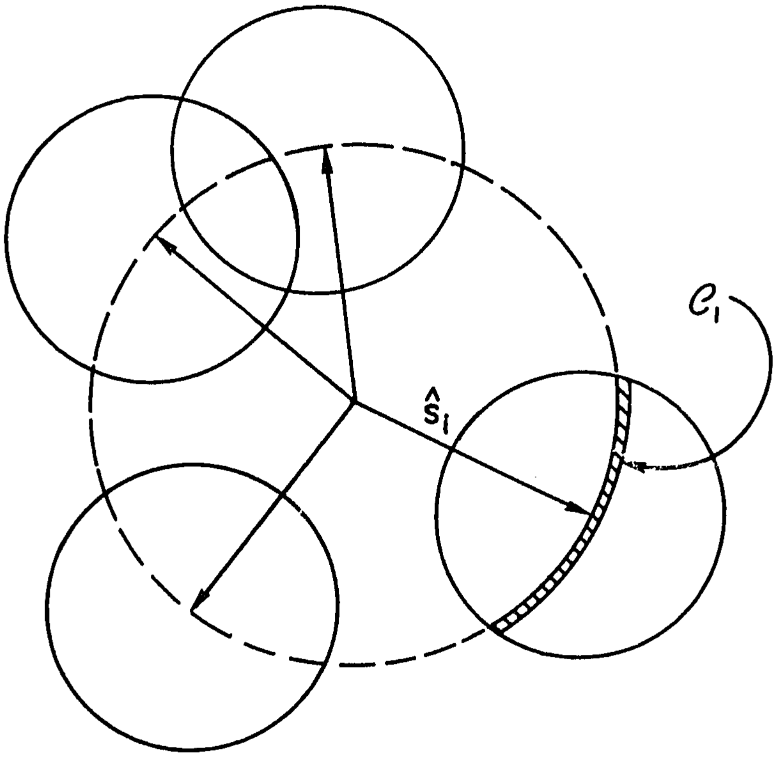

Fig. 1.

every fixed r, and, again without loss of generality, it is sufficient to show that for any arbitrary angle φ, $0 \le \varphi \le \pi$,

$$\sum_i \int_{\mathcal{R}_i \cap \mathcal{C}_i} dS \tag{I.14}$$

is maximized by the regular simplex. Farber also points out that by virtue of the uniform measure assigned by the caps, (I.14) is equivalent to

$$\int_{\mathfrak{D}} dS \tag{I.15}$$

where $\mathfrak{D} = \bigcup_i \mathcal{C}_i$. This formulation removes the dependence on the decision regions.

Chapter II

A SIMPLIFIED PROOF OF LANDAU AND SLEPIAN'S INEQUALITY

The heart of the proof of the simplex conjecture proposed by Landau and Slepian is a geometric inequality which is the extension of a theorem by Fejes-Toth [12]. In this chapter we reexamine this theorem and provide a greatly simplified proof which avoids the complication of spherical geometry and places the problem in a more natural setting. Unfortunately, this new perspective merely allows us to see more clearly why the Landau and Slepian result falls short of the simplex conjecture itself.

The theorem with which we are concerned is the following:

<u>Theorem 1.</u> Consider a spherical cap about vector P_o,

$$\mathcal{C} = \{\hat{X}: \hat{X} \cdot \hat{P}_o > \cos \varphi\}, \tag{II.1}$$

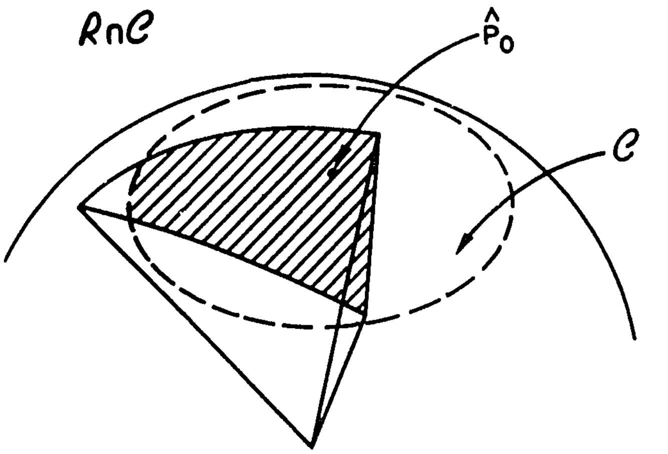

Fig. 2.

which intersects a spherical simplicial cone given by

$$\mathcal{R} = \{X: |X-\hat{P}_o| \le |X-\hat{X}_i| \quad i = 1, \ldots, k\} \tag{II.2}$$

To ensure that each boundary of $\mathfrak{R}$ intersects $\mathcal{C}$ we assume $\hat{X}_i \cdot \hat{P}_o > \cos 2\varphi$. Then if $g_r(r)$ is a monotonically decreasing function of r

$$\int_{\mathfrak{R}\cap\mathcal{C}} g_r(|\hat{P}_o-\hat{X}|)dS \le \int_{\mathcal{C}} g_r(|\hat{P}_o-\hat{X})|dS - k\int_{\mathcal{W}} g_r(|\hat{P}_o-\hat{X}|)dS \qquad \text{(II.3)}$$

whenever

$$\int_{\mathfrak{R}\cap\mathcal{C}} dS = \int_{\mathcal{C}} dS - k\int_{\mathcal{W}} dS \qquad \text{(II.4)}$$

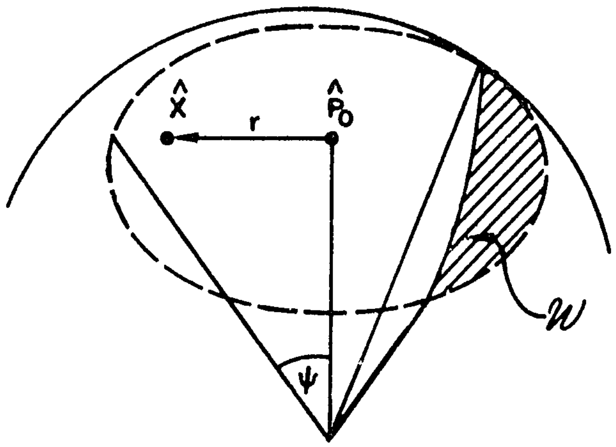

Fig. 3.

where dS is a differential element of surface of the unit hypersphere and $\mathcal{W}$ is a region of the cap cut off by a single hyperplane. If $g_r(r)$ is strictly monotonically decreasing over the entire cap, then equality holds in (II.3) if and only if $\mathcal{C}\cap\mathfrak{R}^c$ is the union of k non-intersecting regions with the form of $\mathcal{W}$.

The Landau and Slepian proof ([5], pp. 1262-1272) is in two parts. In the first they examine the two functions

$$h_r(w) = \int_{\mathcal{W}} g_r(|P_o-X|)dS \qquad \text{and} \qquad w = \int_{\mathcal{W}} dS \qquad \text{(II.5)}$$

as the size of $\mathcal{W}$ varies for an arbitrary but fixed size cap.

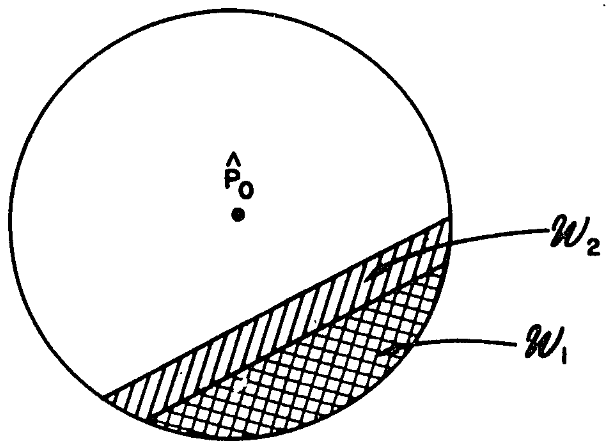

Fig. 4.

Evaluating the integrals above they arrive at the equation

$$\frac{dh_r(w)}{dw} = \int_a^b l_b(x)\ \hat{g}(x)\,dx$$

where

$$l_b(x) = \frac{\left(1 - \frac{x}{b}\right)^{\frac{n-4}{2}}}{\frac{2b}{n-2}\left(1-\frac{a}{b}\right)^{\frac{n-2}{2}}} \qquad a \le x \le b$$

Using a result reprinted in Appendix A and the form of $l_b(x)$ for $n = 3$, $n = 4$, and $n > 4$, they show that this implies

$$w_2 > w_1 \Rightarrow \frac{dh_r(w_2)}{dw} > \frac{dh_r(w_1)}{dw} \qquad \text{(II.6)}$$

which means that the function $h_r(w)$ is convex. Thus if single hyper-planes cut off a total surface W_o from $\mathcal{C}$, the least probability $g_r(X)$ will be removed if each hyperplane cuts off the same amount.

In the second part of the proof they compare a region $\mathcal{V}$, which is the intersection of the cap with a convex cone, with a region $\mathcal{W}$ cut off from the cap by a single hyperplane. Assuming the size of $\mathcal{W}$ is adjusted so that

$$w \equiv \int_{\mathcal{W}} dS = v \equiv \int_{\mathcal{V}} dS \qquad \text{(II.7)}$$

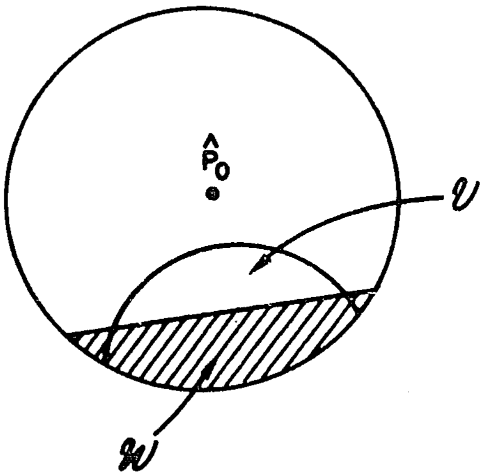

Fig. 5.

they show that

$$I_{\mathcal{V}} \equiv \int_{\mathcal{V}} g_r(|\hat{P}_o - X|)\,dS \geq \int_{\mathcal{W}} g_r(|\hat{P}_o - X|)\,dS \equiv I_{\mathcal{W}} \qquad \text{(II.8)}$$

Their technique is to define two functions

$$w(\Theta) = (n-2) - \text{dimensional content of } \mathcal{B}_{\mathcal{W}}(\Theta)$$

$$\text{(II.9)}$$

$$v(\Theta) = (n-2) - \text{dimensional content of } \mathcal{B}_{\mathcal{V}}(\Theta)$$

where

$$\mathcal{B}_{\mathcal{W}}(\Theta) = \{X:\ X \cdot \hat{P}_o = \cos\Theta,\ X \in \mathcal{W}\}$$

$$\text{(II.10)}$$

and

$$\mathcal{B}_{\mathcal{V}}(\Theta) = \{X:\ X \cdot \hat{P}_o = \cos\Theta,\ X \in \mathcal{V}\}$$

so that

$$v = \int_o^{\varphi} v(\Theta)\,d\Theta = w = \int_o^{\varphi} w(\Theta)\,d\Theta$$

and

$$I_{\mathcal{W}} = \int_o^{\varphi} g_r(\sqrt{2-2\cos\Theta})\ w(\Theta)\,d\Theta$$

$$I_{\mathcal{V}} = \int_o^{\varphi} g_r(\sqrt{2-2\cos\Theta})\ v(\Theta)\,d\Theta.$$

They then prove that

$$v(\Theta) \ge w(\Theta) \ , \quad 0 \le \Theta \le \varphi' \qquad \text{(II.11)}$$

$$v(\Theta) \le w(\Theta) \ , \quad \varphi' \le \Theta \le \varphi$$

by showing that

$$v(\Theta) = w(\Theta) \Rightarrow \frac{dw(\Theta)}{d\Theta} \ge \frac{dv(\Theta)}{d\Theta} \ . \qquad \text{(II.12)}$$

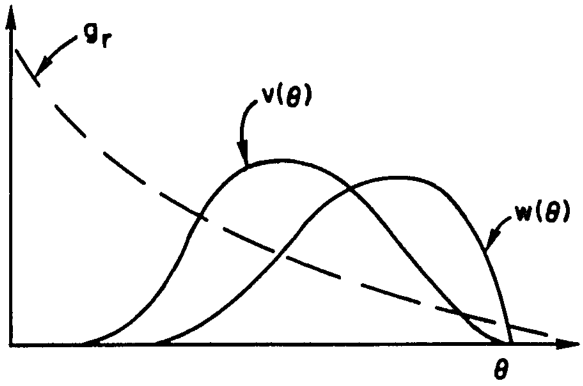

Fig. 6.

The inequality $I_v \ge I_w$ follows from Appendix A, which simply gives mathematical rigor to the obvious idea that $v(\Theta)$ concentrates more of its mass in areas where g_r is large than does $w(\Theta)$ (see Fig. 6).

To demonstrate implication (II.12) the authors construct three associated spherically convex figures $\Gamma_v(Q)$, $\Gamma_w(Q)$, and $\Gamma_w(Q^*)$ for which the functions $\overline{v}(\Theta)$, $\overline{w}(\Theta)$, and $\overline{w}^*(\Theta)$ respectively have the property

$$\frac{dv(\Theta)}{d\Theta} \le \frac{d\overline{v}}{d\Theta}(\Theta) \qquad \frac{d\overline{w}(\Theta)}{d\Theta} \le \frac{dw^*(\Theta)}{d\Theta} = \frac{dw}{d\Theta}(\Theta) \qquad \text{(II.13)}$$

whenever

$$v(\Theta) = \overline{v}(\Theta) = \overline{w}(\Theta) = \overline{w}^*(\Theta) = w(\Theta) \ .$$

An expression for $\overline{v}(\Theta)$ is then developed which enables them to conclude

$$\frac{d\overline{v}}{d\Theta}(\Theta) \le \frac{d\overline{w}(\Theta)}{d\Theta} \qquad \text{(II.14)}$$

The details of the proof involve complicated expressions in spherical coordinates too long to be included here.

As a first step in our simplified proof, we note that the spherical coordinates are an unnecessary bother and serve only to obscure the workings of the problem. Since we can restrict our attention to caps of the form $X \cdot \hat{P}_o \geq 0$, we can perform a spherical projection of the problem on to the $(n-1)$-flat $X \cdot \hat{P}_o = 1$ by letting

$$\hat{X} \rightarrow l(\hat{X}) = \frac{\hat{X}}{\hat{X} \cdot \hat{P}_o} \quad \text{for} \quad \hat{X} \in \mathfrak{R} \tag{II.15}$$

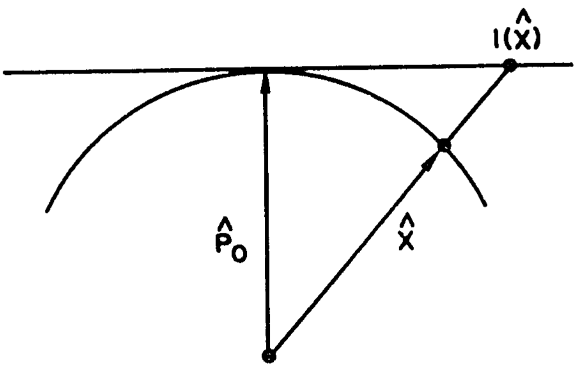

Fig. 7

Now we look at the problem in the $(n-1)$-dimensional space of this $(n-1)$-flat where the tip of $\hat{P}_o$ is the new origin. The $(n-1)$-flat boundaries of $\mathfrak{R}$ are projected into $(n-2)$-flats; the cap is projected into a hypersphere $\mathcal{C}^o$.[1] Letting Y be a generic vector in the $(n-1)$-space $\{Y: Y = X - \hat{P}_o, \ X \cdot \hat{P}_o = 1\}$, we can define a function $s(|Y|)$ such that

$$\int_{\mathfrak{R}^o \cap \mathcal{C}^o} s(|Y|)dV_{n-1} = \int_{\mathfrak{R} \cap \mathcal{C}} dS_n \tag{II.16}$$

1. We shall hereafter use the superscript zero to denote a spherically projected quantity.

In other words, $s(|Y|)$ is the n-dimensional solid angle subtended by a differential volume element in the (n-1)-flat. Similarly we can define a projected density $g_r^o(|Y|)$ such that

$$\int_{\mathfrak{R}^o \cap \mathcal{C}^o} g_r^o(|Y|)dV_{n-1} = \int_{\mathfrak{R}\cap\mathcal{C}} g_r(|\hat{X}-\hat{P}_o|)dS_n \qquad \text{(II.17)}$$

It is easy to see that $s(r_{n-1})$, $g_r^o(r_{n-1})$, and

$$m(r_{n-1}) = g_r^o(r_{n-1})/s(r_{n-1}) \qquad \text{(II.18)}$$

are all monotonically decreasing functions. The problem now becomes one of minimizing

$$\int_{\mathfrak{R}^{o^c} \cap \mathcal{C}^o} g_r^o(r)dV_{n-1} \quad \text{given} \quad \int_{\mathfrak{R}^{o^c} \cap \mathcal{C}^o} s(r)dV_{n-1} = \text{Const.} \qquad \text{(II.19)}$$

In a manner similar to that of Landau and Slepian we can write

$$\mathfrak{R}^{o^c} \cap \mathcal{C}^o = \{Y \in \mathcal{C}^o:\ Y \cdot Y_i \geq 1 \quad \text{for some} \quad i = 1,\ldots, k\} \qquad \text{(II.20)}$$

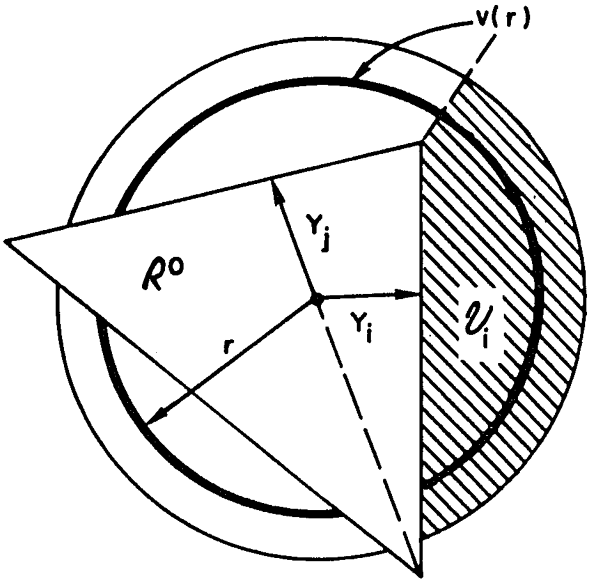

Fig. 8.

$\mathfrak{R}^{o^c} \cap \mathcal{C}^o$ can be decomposed into at most k convex regions $\mathcal{V}_i$ where

$$\mathcal{V}_i = \{Y \in \mathcal{C}^o \colon \; Y \cdot Y_i > Y \cdot Y_j \geq 1 \quad \text{for all} \quad j \neq i\} \qquad \text{(II.21)}$$

and we neglect to assign a set of measure zero. The definition has been auspiciously chosen so that

$$Y \in \mathcal{V}_i, \quad aY \in \mathcal{C}^o \Rightarrow aY \in \mathcal{V}_i \quad \text{for} \quad a \geq 1$$

since (II.22)

$$Y \cdot Y_i > Y \cdot Y_j \geq 1 \Rightarrow aY \cdot Y_i > aY \cdot Y_j \geq 1 \; .$$

This fact will facilitate a later proof.

The astute reader may have already guessed our ploy: if we can show that the set of $\mathcal{V}_i$'s can be replaced by a set of pieces of hypersphere $\mathcal{C}^o$ of the same size, each cut off by a single hyperplane, then projecting back down onto the cap $\mathcal{C}$ will yield the desired result. The next step is to formulate the correct characterization of the region $\mathfrak{R}^{o^c} \cap \mathcal{C}^o$. Since

$$\int_{\mathfrak{R}^{o^c} \cap \mathcal{C}^o} g_r(r)\, dV_{n-1} = \int_0^\infty m(r) s(r) v(r)\, dr$$

where $v(r)$ is the $(n-2)$-dimensional content of the intersection of a shell $|Y| = r$ with $\mathcal{C}^o$ and $m(r)$ and $s(r)$ are as previously defined, integration by parts gives

$$\int_0^\infty m(r)s(r)v(r)dr = m(\infty)\int_0^\infty s(r)v(r)dr - \int_0^\infty\int_0^r s(r')v(r')dr' \frac{d}{dr}m(r)\, dr$$

$$= m(\infty)\int_0^\infty s(r)v(r)dr + \int_0^\infty\int_0^r s(r')v(r')dr'\left(-\frac{d}{dr}\, m(r)\right) dr \qquad \text{(II.23)}$$

The term in parentheses is positive and so we can see that if

$$\int_0^\infty s(r)v(r)dr = \int_0^\infty s(r)w(r)dr \qquad \text{(II.24)}$$

and

$$\int_0^r s(r)v(r)dr \geq \int_0^r s(r)w(r)dr \quad \text{for all } r>0 \qquad \text{(II.25)}$$

then

$$\int_0^\infty g_r^o(r)v(r)dr \geq \int_0^\infty g_r^o(r)w(r)dr \qquad \text{(II.26)}$$

Eventually we will arrive at the fortunate conclusion that the w(r) corresponding to a collection of single hyperplane pieces does just that.

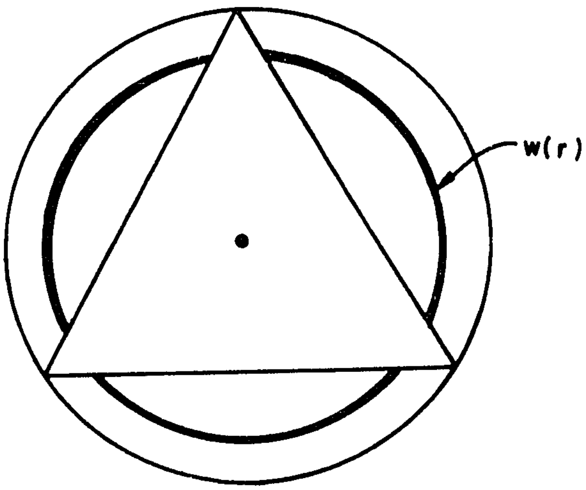

Fig. 9.

The question of what property $w(r)$ must have to achieve this inequality is answered by the following lemma.

Lemma 1: If for any $r > 0$

$$V(r) \equiv \int_0^r v(r)dr < \int_0^r w(r)dr \equiv W(r) \Rightarrow v(r) < w(r) \tag{II.28}$$

and

$$\int_0^\infty s(r)v(r)dr = \int_0^\infty s(r)w(r)dr \ , \tag{II.29}$$

then

$$\int_0^{r_1} s(r)v(r)dr \geq \int_0^{r_1} s(r)w(r)dr \quad \text{for all} \quad r_1 > 0 \ . \tag{II.30}$$

Proof. $(V(r) - W(r) < 0 \Rightarrow \quad v(r) - w(r) < 0)$

means that

$$(V(r_o) - W(r_o) < 0) \Rightarrow (V(r) - W(r) < 0 \quad \text{for} \quad r \geq r_o) \tag{II.31}$$

since

$$\frac{d}{dr}(V(r) - W(r)) = v(r) - w(r) < 0 \ .$$

Thus we can find an r_o such that

$$V(r) - W(r) \geq 0 \qquad \text{for} \quad r \leq r_o \tag{II.32}$$

and

$$V(r) - W(r) < 0 \qquad \text{and} \quad v(r) - s(r) < 0 \qquad \text{for} \quad r > r_o \tag{II.33}$$

Integrating (II.30) by parts gives

$$\int_0^{r_1} (v(r)-w(r))s(r)dr$$

$$= \int_0^{r_1} (v(r)-w(r))dr\ s(r_1) + \int_0^{r_1}\int_0^{r} (v(r')-w(r'))dr'\left(-\frac{d}{dr}s(r)\right)dr$$

$$= (V(r_1)-W(r_1))s(r_1) + \int_0^{r_1} (V(r)-W(r))\left(-\frac{d}{dr}s(r)\right)dr \tag{II.34}$$

Suppose that

$$\int_0^{r_1} (v(r)-w(r))s(r)dr < 0$$

By (II.32) and (II.34) this is possible only if $r_1 > r_o$. But

$$\int_0^{\infty}(v(r)-w(r))s(r)dr = \int_0^{r_1} (v(r)-w(r))s(r)dr + \int_{r_1}^{\infty}(v(r)-w(r))s(r)dr$$

The second part of (II.33) implies that the integrand of the second integral is negative. Consequently the right hand side is less than zero, contradicting our hypothesis. ∎

To employ three-dimensional terminology, we have left to show that when k $\mathcal{W}$-regions cut off as much volume from a hypersphere of arbitrary size as does $\mathcal{V}$, they cut off at least as much surface area. Examine region $\mathcal{V}_i$ which cuts off a volume

$$V_i = \int_0^{r} v_i(r)dr \tag{II.35}$$

from a hypersphere of radius r. We can assume $\mathcal{V}_i$ does not contain the origin. We construct an associated single hyperplane region $\mathcal{W}_i$ so that

$$W_i = \int_0^r w_i(r)dr = V_i \ . \tag{II.36}$$

We will then compare $v_i(r)$ and $w_i(r)$. As a preliminary exercise, however, we prove a lemma concerning $W_i(r)$ and $w_i(r)$. Its object is to show that the ratio of surface to volume of a hyperplane slice of a hypersphere is smallest near the center of the hypersphere. In other words, we are showing that the end slice of an n-dimensional watermelon has the most rind.

Lemma 2. Given an n-dimensional hypersphere $\mathcal{B}(r)$ intersected by a single hyperplane $Y \cdot \hat{Y}_o = t$ at distance t from the origin, denote the n-dimensional content of the set $\{Y: Y \in \mathcal{B}(r),\ Y \cdot \hat{Y}_o > t\}$ by $W_o(t)$ and the $(n-1)$-dimensional content of the set $\{Y: |Y| = r,\ Y \cdot \hat{Y}_o > t\}$ by $w_o(t)$.

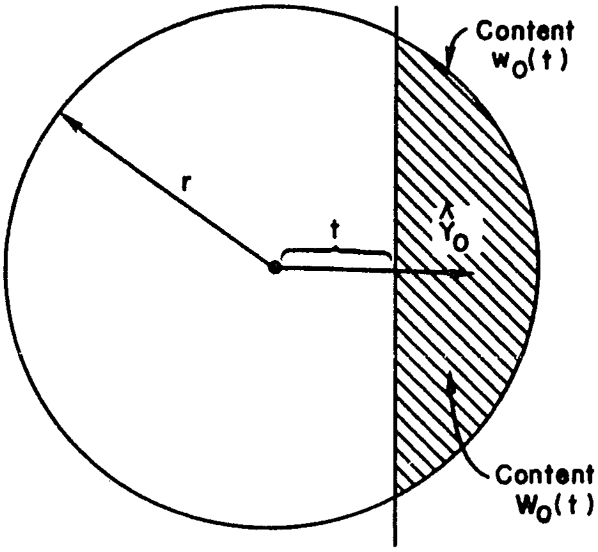

Fig. 10

$$\frac{\frac{dw_o(t)}{dt}}{\frac{dW_o(t)}{dt}} = \frac{dw_o(t)}{dW_o(t)}$$

is a monotonically increasing function of t, and consequently a monotonically decreasing function of $W_o(t)$ for all $n > 1$. Thus $w_o(t)$ is a concave function of $W_o(t)$ (Fig. 14).

<u>Proof.</u>

$$\frac{dW_o(t)}{dt} = - k_n r_1^{n-1}$$

where k_n = volume of a unit $(n-1)$-sphere.

$$\frac{dw_o(t)}{dt} = - (n-1) k_n r_1^{n-2} \cdot \frac{r}{r_1}$$

$$\frac{\frac{dw_o(t)}{dt}}{\frac{dW_o(t)}{dt}} = (n-1) \frac{r}{r_1^2} = (n-1) \frac{r}{(r^2-t^2)} \qquad \blacksquare \quad \text{(II.37)}$$

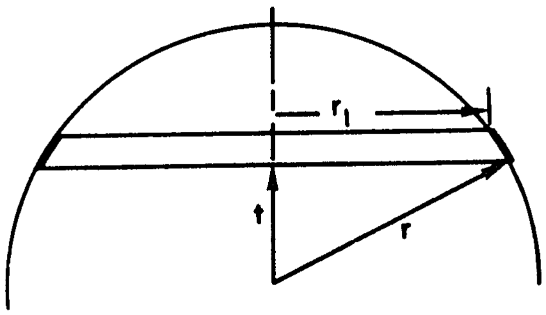

Fig. 11.

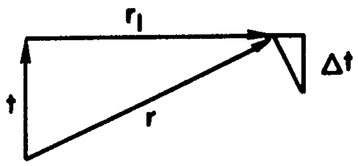

Fig. 12.

We can now give the last required lemma relating the content functions of the regions we are comparing.

Lemma 3. $$V(r) = W(r) \Rightarrow v(r) \le w(r) \tag{II.38}$$

Proof. In what follows, we fix r and consider that

$$\mathcal{V} = \bigcup_{i=1}^{k} \mathcal{V}_i \tag{II.39}$$

Since $\mathcal{V}_i \cap \mathcal{V}_j = \phi$, $i \neq j$, the corresponding content functions satisfy

$$V(r) = \sum_{i=1}^{k} V_i \tag{II.40}$$

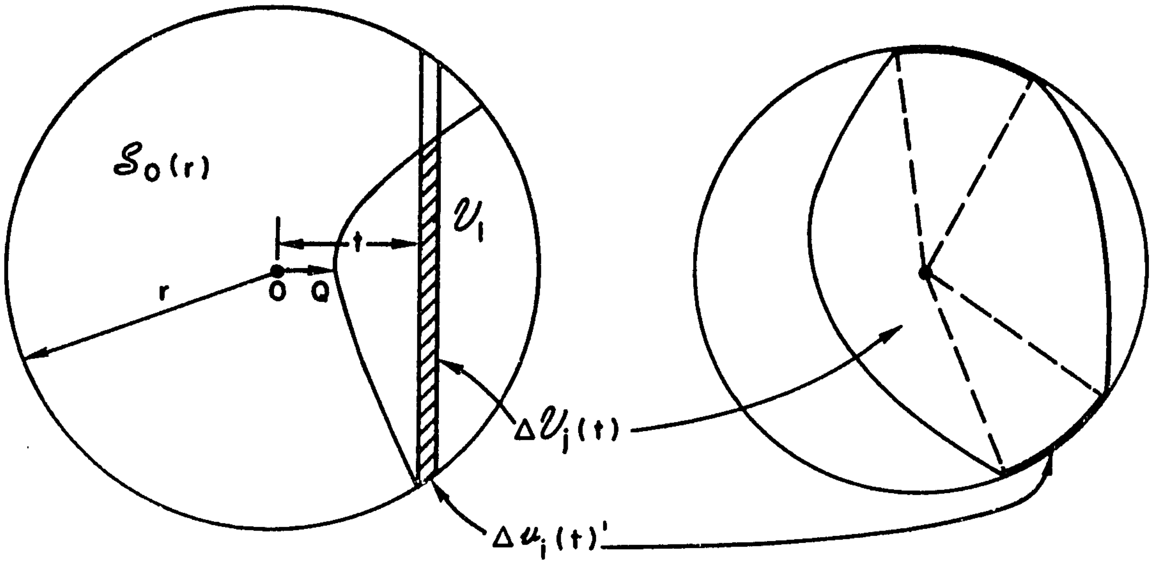

Fig. 13.

Let $\hat{Y}_o$ of lemma 2 be $\frac{OQ}{|OQ|}$ where Q is the point of $\mathcal{V}_i$ closest to the origin (Fig. 13). For $t \ge |OQ|$ let $\Delta\mathcal{V}_i(t)$ be the intersection of the hyperplane

$$\{Y:\ Y\cdot\hat{Y}_o = t\} = \mathcal{H}(t) \tag{II.41}$$

with $\mathcal{V}_i$ and $\Delta V_i(t)$ its (n-2)-dimensional content. Likewise let $\Delta\nu_i(t)$ be the intersection $\mathcal{V}_i \cap \mathcal{H}(t) \cap \mathcal{S}_o(r)$ and $\Delta v_i(t)$ its (n-3)-dimensional content. $\mathcal{V}_i(t)$ is convex, being the intersection of two convex regions, and further, by (II.22) and the definition of $\hat{Y}_o$, $\Delta\mathcal{V}_i(t)$ contains the origin of the (n-3)-dimensional hypersphere

$$\mathcal{H}(t) \cap \mathcal{B}(r) \equiv \Delta\mathcal{B}(t) \tag{II.42}$$

for all $t \geq |OQ|$. It follows immediately that

$$\frac{v_i(t)}{V_i(t)} \leq \frac{\text{surface of } \Delta\mathcal{B}(t)}{\text{volume of } \Delta\mathcal{B}(t)} = \frac{\Delta w_o(t)}{\Delta W_o(t)} = \frac{dw_o(t)}{dW_o(t)} \tag{II.43}$$

since $\Delta\mathcal{V}_i(t)$ must contain the cones with base $\Delta\nu_i(t)$ and vertex at the origin. Consequently

$$v_i = \int_{|OQ|}^{r} \frac{\Delta v_i(t)}{\Delta V_i(t)} \Delta V_i(t)dt \leq \int_{|OQ|}^{r} \frac{\Delta w_o(t)}{\Delta W_o(t)} \Delta V_i(t)dt$$

$$\leq \int_{|OQ|}^{r} \frac{dw_o(t)}{dW_o(t)} \Delta W_i(t)dt = W_i \tag{II.44}$$

where $\Delta W_i(t)$ is the corresponding function of $\mathcal{W}_i$, $(W_i = V_i)$, and the last inequality follows from Appendix A.

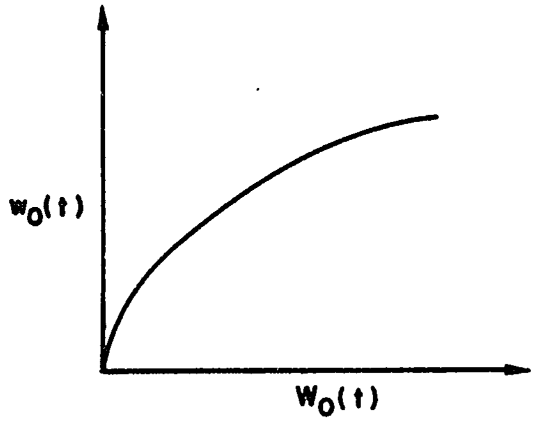

Fig. 14.

Putting pieces together, we have

$$V(r) = \sum_{i=1}^{k} V_i = \sum_{i=1}^{k} W_i = kW_o(t^*) = W(r)$$

$$\Rightarrow \quad v(r) = \sum_{i=1}^{k} v_i \le \sum_{i=1}^{k} w_i \le kw_o(t^*) = w(r) \tag{II.45}$$

The last inequality is an application of the concavity proved in Lemma 2, with t^* chosen to make $\sum_{i=1}^{k} W_i = kW_o(t^*)$. ∎

With all the work out of the way, we summarize by giving the completed proof.

Proof of Theorem 1. By the remarks preceding (II.19), it is sufficient to solve the projected problem of minimizing

$$\int_{\mathfrak{R}^o \cap \mathcal{C}^{o^c}} g_r^o(r)\,dV_{n-1} \quad \text{given} \quad \int_{\mathfrak{R}^o \cap \mathcal{C}^{o^c}} s(r)\,dV_{n-1} = S_i \quad \text{(a constant).} \tag{II.46}$$

Defining $v(r)$ corresponding to $\mathcal{V} = \mathcal{R}^{o^c}$, Eqs. (II.23) to (II.26) indicate that it is sufficient to minimize

$$\int_o^r s(r)v(r)dr \quad \text{given} \quad \int_o^\infty s(r)v(r)dr = S_1 \tag{II.47}$$

Now we construct a region $\mathcal{W}^o$ using a single hyperplane intersecting $\mathcal{C}^o$ such that

$$k\int_{\mathcal{W}^o} s(r)dV_{n-1} = S_1 \tag{II.48}$$

and let

$$w(r) = kw^o(r) \ . \tag{II.49}$$

By Lemma 3, we know

$$V(r) = W(r) \Rightarrow v(r) \le w(r) \tag{II.50}$$

so unless $\mathcal{V}$ does not intersect a hypersphere of radius r at all, we have

$$V(r) < W(r) \Rightarrow v(r) < w(r) \ . \tag{II.51}$$

The condition for Lemma 1 being satisfied, this implies

$$\int_o^r s(r)w(r) \le \int_o^r s(r)v(r)dr \tag{II.52}$$

Since $\mathcal{W}^o$ must be the projection of $\mathcal{W}$, we then have

$$\int_{\mathcal{R}^c \cap \mathcal{C}} g_r(|P_o-X|)dS \ge k\int_{\mathcal{W}} g_r(|P_o-X|)dS \tag{II.53}$$

from which the required result follows. The conditions for equality arising in the use of Appendix A and Lemma 1 can be seen to imply strict inequality unless $g_r(r)$ is constant over $\mathcal{R}^c \cap \mathcal{C}$ or $\mathcal{R}^c \cap \mathcal{C}$ is itself the union of k regions with the form $\mathcal{W}$. ∎

Having obtained the inequality of Theorem 1, we naturally inquire how much it is worth. Sadly the reply is not a great deal. Consider the projection of a four-dimensional cap $(n=4)$ and an intersecting spherical simplex into three-dimensional space. In order to apply Theorem 1 we have to be able to construct a region which cuts off pieces of the sphere with only one plane when the intersection has an arbitrary content. Landau and Slepian assumed that a regular tetrahedron would do the job. If we observe what happens as we grow a sphere inside the tetrahedron, however, we see that first circles appear in the faces, then the circles hit the edges, and finally the sphere will contain the tetrahedron. The assumption amounts to saying that at the point where the circles hit the edges, the tetrahedron is contained within the sphere. Chapter I of Farber's thesis is devoted to showing the existence of holes around the vertices at this point for all $n > 3$.

Ironically, although Farber states that Theorem 1 can be used to show optimality over a range θ_o to $\hat{\theta}$, where the parts of the <u>non-zero $g_r(r)$</u> region are cut off by single hyperplanes, such is not the case. To apply the theorem, there must exist a collection of $n+1$ caps covering the entire unit hypersphere such that these covering caps, not the non-zero $g_r(r)$ regions, are intersected by single hyperplanes if the signal set is a regular simplex. For $n > 3$ this is not possible. To put it another way, if $g_r^o(r)$ is zero for $r \geq r_1$, we can write the contribution to the probability of correct decoding of $\mathcal{R}^o$ as

$$p_c \quad \text{due to} \quad \mathfrak{R}^o = \int_o^{r_1} g_r^o(r) r^o(r) dr \tag{II.54}$$

where $r^o(r)$ is the function corresponding to $\mathfrak{R}^o$. Also, the solid angle at the origin devoted to $\mathfrak{R}^o$ is

$$\int_o^{\infty} s(r) r^o(r) dr \tag{II.55}$$

Theorem 1 tells us that for

$$\int_o^{r_1} s(r) r^o(r) dr = \text{constant} \tag{II.56}$$

we maximize (II.54) if $\mathfrak{R}^o$ is a regular simplex. But what we need to hold constant is (II.55), and (II.56) will yield (II.55) only if

$$\int_{r_1}^{\infty} s(r) r^o(r) dr \tag{II.57}$$

is minimized by the regular simplex. Even restricting our attention to $\mathfrak{R}^o$'s formed by equidistant hyperplanes, (II.57) turns out to be roughly the same as Conjecture 3 of the following chapter, which is sufficient to prove the simplex conjecture itself.

Chapter III.

SOME NEW FORMULATIONS

Of the various techniques that could be used in a proof of the conjecture, mathematical induction seems to be one of the most natural. The following hierarchy of conjectures on the regular simplex suggests some of the possible properties that could be used as basis for such a proof.

A. Conjecture 1: Minimum dS/dr for fixed S and r.

In the second part of Landau and Slepian's proof of Theorem 1, they consider two functions $v(\Theta)$ and $w(\Theta)$ (see Eq. (II.9)) corresponding to two competing regions cutting off pieces of a cap. They show

$$\begin{aligned} v(\Theta) &\gtrsim w(\Theta) \qquad \Theta < \Theta' \\ v(\Theta) &\leq w(\Theta) \qquad \Theta \geq \Theta' \end{aligned} \tag{III.1}$$

which means that $\mathcal{V}$ cuts off more probability per cap content than does $\mathcal{W}$. The same idea could be incorporated in a proof addressed to the decoding regions themselves, rather than these complement regions. If $r_1(\Theta)$ and $r_2(\Theta)$ correspond to an arbitrary region and a symmetric region respectively such that

$$\int_0^{\pi/2} (r_1(\Theta) - r_2(\Theta))d\Theta = 0 \tag{III.2}$$

then we might like to show

$$\begin{aligned} r_2(\Theta) &\geq r_1(\Theta) \qquad \Theta < \Theta' \\ r_2(\Theta) &\leq r_1(\Theta) \qquad \Theta \geq \Theta' \end{aligned} \tag{III.3}$$

Thus $r_2(\Theta)$ would concentrate more mass in high probability areas and give better performance. It would be sufficient to show that

$$r_2(\Theta) = r_1(\Theta) \Rightarrow \frac{dr_2}{d\Theta} \le \frac{dr_1}{d\Theta} \qquad \text{(III.4)}$$

But the projection of Chapter I can be employed to simplify matters, since this is equivalent to

$$r_2^o(r) = r_1^o(r) \Rightarrow \frac{dr_2^o}{dr} \le \frac{dr_1^o}{dr} \qquad \text{(III.5)}$$

where r_2^o and r_1^o are the related projected functions. This latter statement amounts to showing when an arbitrary collection of $n+1$ or fewer hyperplanes contains the same shell surface of a shell $|X| = r$ as does a symmetric configuration, the symmetric contains less of the shell $|X| = r + \Delta r$. Our first conjecture makes this assertion.

Before giving the precise statement of the conjecture, we briefly reprove the property of the complement regions demonstrated by Landau and Slepian in their Appendix C. Our method of proof enables us to expose some of the pitfalls encountered in attempting to correct the shortcomings of the Fejes-Toth theorem.

As in Chapter II, let $\mathcal{V}_i$ be a convex region intersecting an n-dimensional hypersphere, and define $v_i(r)$ to be the (n-1)-dimensional content of the intersection of a shell $\mathcal{S}_o(r) = \{X : |X| = r\}$ with $\mathcal{V}_i$. Similarly define $w(r)$ corresponding to a region $\mathcal{W}$ which is cut off from the hypersphere by a single hyperplane. If we can show

$$w(r) = v_i(r) \Rightarrow \frac{dw(r)}{dr} > \frac{dv_i(r)}{dr} ,$$

then this implies a single crossing of $w(r)$ and $v(r)$, and thus of $v(\Theta)$ and $w(\Theta)$ of (II.9). Again let Q be the point in $\mathcal{V}_i$ closest to the origin. Define $\mathfrak{D}$ to be the cone with base $\upsilon_1(r+\Delta r) = \mathcal{S}_o(r+\Delta r) \cap \mathcal{V}_i$. Since $\mathcal{V}_i$ is convex, the intersection $\mathfrak{D} \cap \mathcal{S}_o(r)$ is contained in $\upsilon_i(r)$.

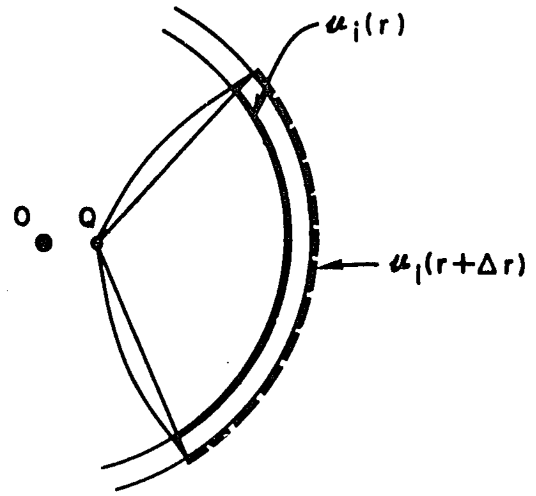

Fig. 15.

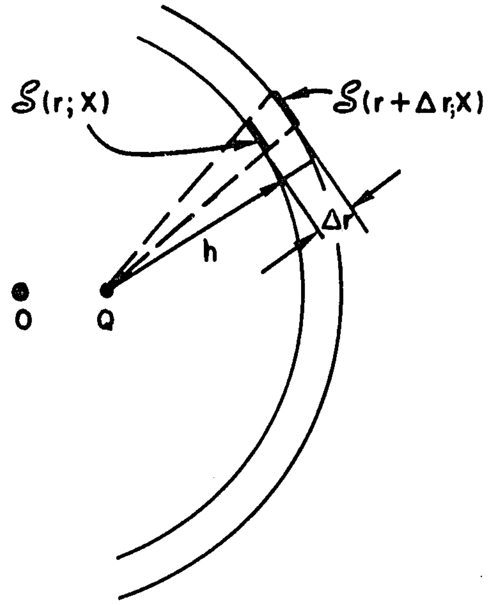

Fig. 16.

Consequently, the content of the projection of $\mathcal{V}_i(r)$ onto the shell $\mathcal{S}_o(r+\Delta r)$ is greater than or equal to the content $\upsilon_i(r+\Delta r)$. The projection ratio

$$\frac{S(r+\Delta r;X)}{S(r;X)} = 1 + \frac{dS/dh}{S(r)}\Delta r \qquad S = kh^{n-1}$$

$$= 1 + \frac{k(n-1)h^{n-2}}{kh^{n-1}}\Delta r \qquad \frac{dS}{dh} = k(n-1)h^{n-2}$$

$$= 1 + (n-1)\frac{1}{h}\Delta r \qquad \text{(III.6)}$$

is a monotonic decreasing function of h, the distance from Q to the hyperplane which touches the shell at the point being projected, X. Since h is a monotonically increasing function of the angle between X and OQ, it is trivial to show that the "shadow" of $\upsilon(r)$ will be maximized if the area of $\upsilon(r)$ is concentrated closest to OQ; that is, in a cap of the form

$$\mathcal{C}_1 = \{X: \frac{X}{|X|} \cdot \frac{OQ}{|OQ|} > \cos \varphi\} . \tag{III.7}$$

The shadow is strictly increased by moving Q out along OQ, so we move Q out until it is in the hyperplane that cuts off $\mathcal{C}_1$. But at this point the shadow exactly equals $w(r+\Delta r)$. Therefore we have

$$w(r) = v(r) \Rightarrow w(r+\Delta r) \geq v(r+\Delta r) \tag{III.8}$$

and thus

$$\frac{dw}{dr} \geq \frac{dv}{dr} .$$

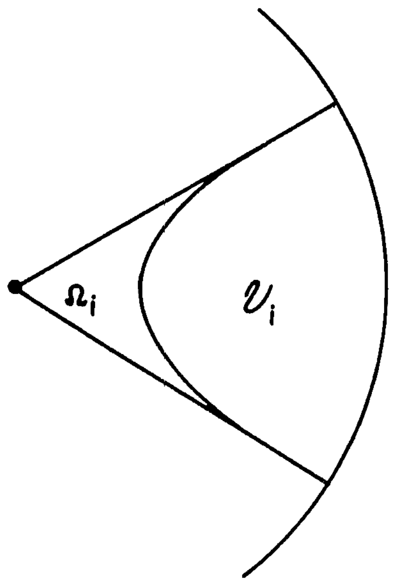

Fig. 17.

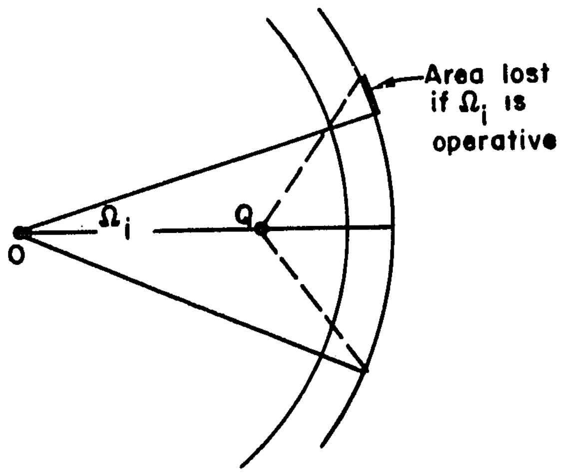

Fig. 18.

The catch is that in higher dimensions there is not enough room to construct caps for all the complementary regions $\mathcal{V}_i$. We have shown that we can improve performance by flattening each $\mathcal{V}_i$ out against the wall of the hypersphere, but unfortunately there is simply not enough room on the wall for all of the $n+1$ $\mathcal{V}_i$'s. Would it not be possible to restrict the n-dimensional angle subtended by $\mathcal{V}_i$, say Ω_i, and demonstrate the same sort of improvement? We regret to say no, or rather, not easily, as the following discussion demonstrates: since we are ultimately trying to prove that for a fixed (n+1)-dimensional angle Ω_{n+1}, we maximize the intersection with a spherical cap by symmetrizing the simplicial cone about the center of the cap while keeping Ω_{n+1} fixed, we can assume this to be true in n-dimensions. In particular we symmetrize the region $\upsilon_i(r)$ in a regular simplicial cone about $\overline{OQ}$, keeping the subtended angle Ω_i fixed. Since we know that we have maximized the intersection with any cap of the type $\mathcal{C}_1$, we can readily prove that the Q-shadow of an area $\upsilon_i(r)$ will be maximized by replacing $\upsilon_i(r)$ with a symmetrized version. In other words, the symmetrized Ω_i is the best replacement for the $\mathcal{C}_1$ under the constraint that the region be bounded by hyperplanes and have n-dimensional solid angle Ω_i. Consequently the following inequalities hold.

Assuming $v(r) = u(r)$,

$$v(r+\Delta r) \leq v_Q(r+\Delta r) \leq u_Q(r+\Delta r) \tag{III.9}$$

where the Q subscript denotes the shadow area and $u(r)$ is the function for a regular simplex cone of size Ω_i. If nature were at all benevolent we would append the equality $u_Q(r+\Delta r) = u(r+\Delta r)$ and the simplex code conjecture would be essentially resolved, but in reality $u(r+\Delta r) < u_Q(r+\Delta r)$ whenever the Ω_i constraint is operative. The shadow is strictly greater whenever the hyperplanes of Ω_i intersect the cap, due to the proximity of Q to the edges (see Fig. 18). The conclusion is that the correction can be made only when it is not needed! Taking advantage of the induction hypothesis is to no avail.

With this elaborate introduction, we present the strongest of our conjectures on the regular simplex. Let

$$\mathcal{H}_i = \{X \in E^n : X \cdot \hat{X}_i \le c_i, |\hat{X}_i| = 1\} \qquad i = 1, \ldots, n+1 \qquad \text{(III.10)}$$

and $\mathcal{J} = \bigcap^{n+1} \mathcal{H}_i$. Let $\tilde{\mathcal{H}}_i$ correspond to a regular simplex, i.e., $\tilde{\hat{X}}_i \cdot \tilde{\hat{X}}_j = \frac{-1}{n}$, $j \neq i$, $c_i = c_j$ for all i, j, and $\tilde{\mathcal{J}} = \bigcap^{n+1} \tilde{\mathcal{H}}_i$. Let $t(r)$ be the (n-1)-dimensional measure of the set $\mathcal{S}_o(r) \cap \mathcal{J}$ and $\tilde{t}(r)$ that of $\mathcal{S}_o(r) \cap \tilde{\mathcal{J}}$, ($\mathcal{S}_o(r) = \{X \in E^n : |X| = r\}$).

<u>Conjecture 1</u>: There exists r_o such that $\tilde{t}(r) \ge t(r)$ for $r \le r_o$, $\tilde{t}(r) \le t(r)$ for $r \ge r_o$.

First we note that the conjecture is definitely false in two dimensions. Consider two concentric rings intersected by two parallel lines at distances c_1 and c_2 from the origin. Assuming both lines intersect both rings, we attempt to minimize the perimeter of the outside ring between the two, $t(r_2)$, while holding the interior perimeter of the inside ring, $t(r_1)$, fixed.

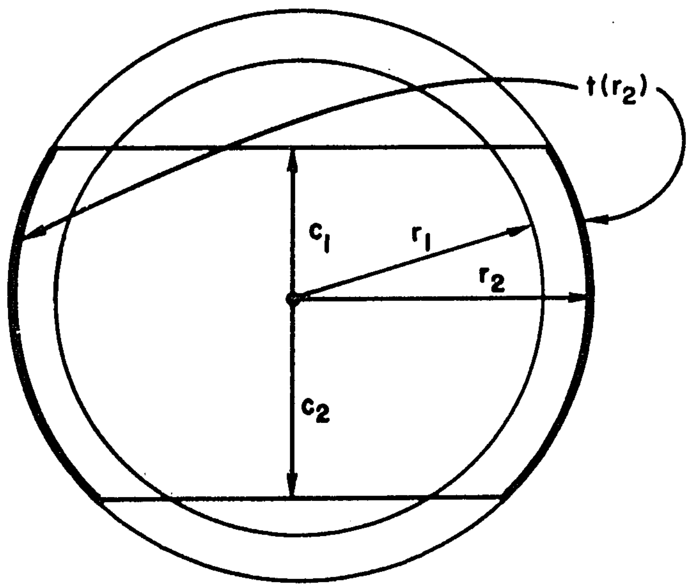

Fig. 19.

$$\frac{dt(r_2)}{dc_1} = -2\frac{d}{dc_1}\sin^{-1}\frac{\left(r_2^2 - c_1^2\right)^{1/2}}{r_2} = 2\,\frac{c_1}{\left(r_2^2 - c_1^2\right)^{-1/2}} \qquad \text{(III.11)}$$

so

$$\frac{dt(r_2)}{dt(r_1)} = \frac{\left(r_1^2 - c_1^2\right)^{1/2}}{\left(r_2^2 - c_1^2\right)^{1/2}}\,, \qquad \text{decreasing in } c_1 \quad .$$

It follows that instead of equalizing c_1 and c_2, we increase the larger until it no longer cuts the inner ring and decrease the smaller accordingly.

This anomaly does not persist in higher dimensions, however. If we examine two n-dimensional shells of radii r_1 and r_2 respectively, we have

$$S_i = \int_c^{r_i} k\left(r_i^2 - c'^2\right)^{\frac{n-2}{2}} \frac{r_i}{\left(r_i^2 - c'^2\right)^{\frac{1}{2}}}\,dc'$$

whence

$$\frac{dS_i}{dc} = -K\left(r_i^2 - c^2\right)^{\frac{n-3}{2}} r_i \qquad \text{(III.12)}$$

so

$$\frac{dS_2}{dS_1} = \frac{\left(r_2^2 - c^2\right)^{\frac{n-3}{2}} r_2}{\left(r_1^2 - c^2\right)^{\frac{n-3}{2}} r_1} = \frac{r_2}{r_1}\left(\frac{r_2^2 - c^2}{r_1^2 - c^2}\right)^{\frac{n-3}{2}} \qquad \text{(III.13)}$$

which is constant for $n = 3$ and monotonically increasing in c for $n > 3$. The necessary convexity for separate caps is thus assured and we may hope that Conjecture 1 will be valid. Nevertheless, the three dimensional situation plagued by the infinite derivative in Eq. (III.11). Consider a sphere inside a tetrahedron with the radius just larger than the distance from the origin to the 1-dimensional edges of the tetrahedron. Looking at one face we see a circle just slightly too large for the triangle which contains it (Fig. 21). We can infer from Eq. (III.12)

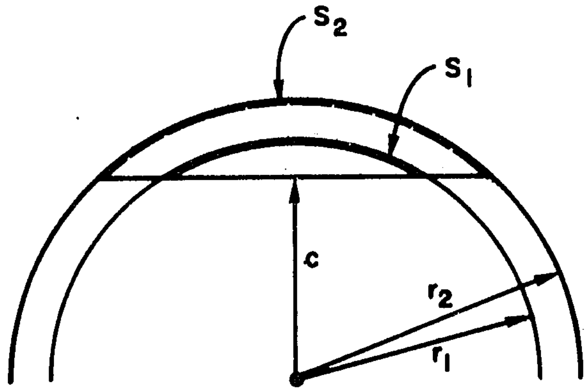

Fig. 20.

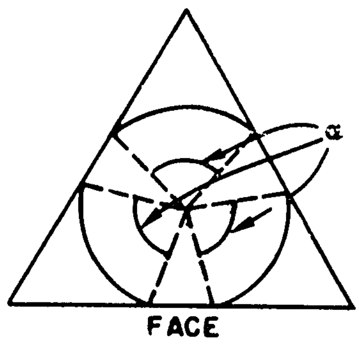

Fig. 21.

that $\frac{dt_3(r)}{dr}$ is proportional to $\alpha \cdot c$, where α is the angle subtended by the interior perimeter and c is the distance from the face to the origin. Now suppose we nudge the face to one side. Because of the infinite derivative we can create a non-zero change in $\frac{dt_3(r)}{dr}$ for an arbitrary small change in $t_3(r)$; Conjecture 1 is suspect in three dimensions also.

In offering this evidence we wish merely to shake the faith of those who believe that the regular simplex must be the extremum for any reasonable problem.

B. Conjecture 2: Minimum Surface Per Volume.

As the next member of the hierarchy, we have, using the notation of Conjecture 1, Conjecture 2:

$$\int_0^{r_o} t(r)dr = \int_0^{r_o} \tilde{t}(r)dr \Rightarrow \tilde{t}(r_o) \le t(r\ ) \quad .$$

The two integrals represent an intersected volume, and so we are hypothesizing that the regular simplex minimizes the intersected surface or shell content for a fixed intersected volume or hypersphere content. As an illustration, we prove Conjecture 2 in two dimensions $(n=2)$:

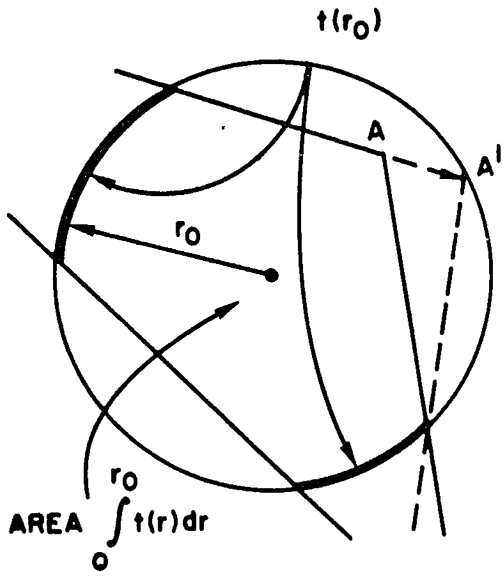

Fig. 22.

Consider three lines intersecting a circle. If two lines form a vertex that is inside the circle, this vertex may be moved out along one of the lines increasing $\int_0^{r_o} t(r)dr$ without changing $t(r_o)$. We can therefore assume that each line cuts off a separate piece. By inspection, the volume cut off as a function of perimeter cut off for a single line is concave, implying that all lines should be the same distance from the center of the circle, which in turn means that the optimum configuration must be equivalent to an equilateral triangle.

This is actually just the specialization of Lemmas 2 and 3 of Chapter II, and when combined with Lemma 1 of that chapter, it provides the simplest known proof for the simplex code conjecture in three dimensions.

To illustrate with one example a recurrent theme we have found in all approaches to the simplex problem we have ever considered, let us begin the construction of an induction proof for Conjecture 2. Assuming it to be true for n-dimensions also that we can eliminate in (n+1)-dimensions all contenders which do not contain the origin, dissect the competing region into n+1 pieces corresponding to each of the faces formed by the H_i,

$$\mathcal{R}_i = \{X: X \cdot \frac{\hat{X}_i}{c_i} \geq X \cdot \frac{\hat{X}_j}{c_j} \quad \text{for all } j, \quad |X| < r_o\} \tag{III.14}$$

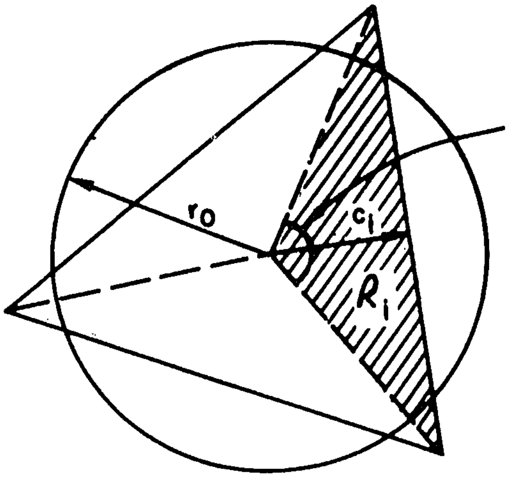

Fig. 23.

This region subtends an (n+1)-dimensional solid angle, say Ω_i, and contains (n+1)-dimensional intersection content $\int_o^{r_o} {}_{n+1}t_{\mathcal{R}_i}(r_{n+1})dr_{n+1}$ and surface ${}_{n+1}t_{\mathcal{R}_i}(r_o)$. We can write

$$\Omega_i = \int_o^\infty {}_nt_{\mathcal{R}_i}(r_n)g_r(r_n)dr_n \tag{III.15}$$

$$ {}_{n+1}t_{\mathcal{R}_i}(r_o) = \int_{(r_o^2-c_i^2)^{1/2}}^\infty {}_nt_{\mathcal{R}_i}(r_n)g_r(r_n)dr_n \tag{III.16}$$

$$\int_0^{r_o} t_{\mathcal{R}_i}(r_{n+1})dr_{n+1} = c_i \int_0^{\left(r_o^2-c_i^2\right)^{1/2}} {}_n t_{\mathcal{R}_i}(r_n)dr_n + \frac{1}{n+1}\, r_o\, {}_{n+1}t_{\mathcal{R}_i}(r_o) \tag{III.17}$$

where ${}_n t_{\mathcal{R}_i}(r_n)$ is the function corresponding to $\mathcal{R}_i \cap \mathcal{H}_i$ in the subspace $\mathcal{H}_i - c_i\hat{X}_i$, the face, and c_i is, of course, the distance from $\mathcal{H}_i$ to the origin. Applying the induction hypothesis, we replace ${}_n t_{\mathcal{R}_i}(r_n)$ by ${}_n\tilde{t}(r_n)$, the function corresponding to a regular simplex face and chosen so that

$$\int_0^{\left(r_o^2-c_i^2\right)^{1/2}} {}_n\tilde{t}(r_n)dr_n = \int_0^{\left(r_o^2-c_i^2\right)^{1/2}} {}_n t_{\mathcal{R}_i}(r_n)dr_n \tag{III.18}$$

We know that ${}_n\tilde{t}(r_n) \le {}_n t_{\mathcal{R}_i}(r_n)$ for $r_n \ge \left(r_o^2-c_i^2\right)^{1/2}$ which implies that

$${}_{n+1}\tilde{t}(r_o) = \int_{\left(r_o^2-c_i^2\right)^{1/2}}^{\infty} {}_n\tilde{t}(r_n)g_r(r_n)dr_n \le t_{\mathcal{R}_i}(r_o) \tag{III.19}$$

If we then increase the size of simplex face until equality holds in (III.19), it is guaranteed that

$$\int_0^{r_o} {}_{n+1}\tilde{t}(r_{n+1})dr_{n+1} = c_i \int_0^{\left(r_o^2-c_i^2\right)^{1/2}} {}_n\tilde{t}(r_n)dr_n + \frac{1}{n+1}\, c_i \cdot {}_{n+1}\tilde{t}(r) \tag{III.20}$$

$$\geq \int_0^{r_o} {}_{n+1}t_{\mathcal{R}_i}(r_{n+1})dr_{n+1}$$

On the other hand, the induction hypothesis also implies, by virtue of Lemma 1 of Chapter II, with ${}_n t_{\mathcal{R}_i}(r_n) = kr^{n-1} - v(r)$ and

${}_n\tilde{t}(r) = kr^{n-1} - w(r)$, that of all (n+1)-dimensional spherical simplexes with at most n hyperplanes sides, the regular simplex maximizes the intersection with any cap. Applying this to the caps formed by $\mathcal{H}_i$, we conclude that symmetrization of Ω_i leads to

$$ {}_{n+1}t^*(r_{n+1}) \le {}_{n+1}t_{\mathcal{R}_i}(r_{n+1}) \quad \text{for } \underline{\text{all}} \quad r_{n+1} > 0 $$

$$ \Rightarrow \int_0^{r_o} {}_{n+1}t^*(r)\,dr_{n+1} \le \int_0^{r_o} {}_{n+1}t_{\mathcal{R}_i}(r_{n+1}) \tag{III.21} $$

so

$$ {}_{n+1}\tilde{t}(r_{n+1}) \ge {}_{n+1}t^*(r_{n+1}) \tag{III.22} $$

with equality only if $\mathcal{R}_i$ is symmetrical or $c_i \ge r_o$.

In short, we can symmetrize maintaining the volume $\int_0^{r_o} {}_{n+1}t_{\mathcal{R}_i}(r_{n+1})dr_{n+1}$ and decreasing ${}_{n+1}t_{\mathcal{R}_i}(r_o)$, but we are guaranteed that the resulting configuration has $\tilde{\Omega}_i > \Omega_i$ in all nontrivial cases where $c_i < r_o$. The desired program to complete the induction is to replace each $\mathcal{R}_i$ with a symmetrized version and then demonstrate a convexity of ${}_{n+1}\tilde{t}(r_o)$ in terms of Ω_i and $\int_0^{r_o} {}_{n+1}\tilde{t}(r_{n+1})dr_{n+1}$, which would allow the equalization of the pieces. Finally, we would put all the pieces together to form a (n+1)-dimensional regular simplex. Clearly this last step is feasible only if

$$ \sum_1^{n+1} \Omega_i' = \sum_1^{n+1} \Omega_i \quad . $$

The violation of the Ω_i constraint has destroyed the program.

Restoring $\int_0^{r_o} {}_{n+1}t_{\mathcal{R}_i}(r_{n+1})dr_{n+1}$ if the Ω_i constraint is respected, necessitates increasing c_i. Under these circumstances, additional information on the comparative forms of ${}_nt_{\mathcal{R}_i}(r_n)$ and ${}_n\tilde{t}(r_n)$ must be extracted to ensure that ${}_{n+1}\tilde{t}(r_{n+1})$ gives better performance on the constrained optimization problem

As a final note on Conjecture 2, we propose the following related conjecture which we also think to be true.

Conjecture 2':

Let $\mathcal{V}$ be a convex region intersecting a hypersphere $\mathcal{B}_n(r)$ such that $\mathcal{V} \cap \mathcal{B}_n(r)$ is contained within a spherical polygon formed by n hyperplanes and subtending a solid angle Ω (as in Fig. 17). Then

$$\int_0^r v(r)dr = \int_0^r w(r)dr \Rightarrow v(r) \leq w(r) \qquad \text{(III.23)}$$

where $v(r)$ and $w(r)$ are the shell intersection functions, and $\mathcal{W}$ is formed by a symmetrical simplicial cone of angle Ω intersecting a single hyperplane.

C. Conjecture 3: Minimum S for all r.

The next conjecture is merely a rephrasing of the last conjecture mentioned in the introduction (Eq. (I.15)). Rather than considering variable size caps on the unit hypersphere, we maintain the fixed angle and vary the radius of the hypersphere. Writing the caps as

$$\mathcal{C}_i = \{X: X \cdot \hat{X}_i > c_o, |X| = r\} \qquad \text{(III.24)}$$

maximizing (I.15) is equivalent to minimizing

$$\int_{\mathcal{J}} ds \tag{III.25}$$

where

$$\mathcal{J} = \{X: X \cdot \hat{X}_i < c_o \quad \text{for all} \quad i, \ |X| = r\} .$$

Using the notation of Conjecture 1, we then have

<u>Conjecture 3.</u>

If $c_i = \tilde{c}_i = c_o$ for all i, so that the $\mathcal{H}_i$ and the $\tilde{\mathcal{H}}_i$ are at the same distance from the origin, then $\tilde{t}(r) \le t(r)$ for all r.

Examine now the collection of $\mathcal{H}_i$'s. Forgetting for a moment the constraint on the distance from each hyperplane to the origin, what change in the intersected shell content, say $S_I(r)$, does a slight movement of $\mathcal{H}_i$ cause?. If the movement of $\mathcal{H}_i$ does not alter the normal to $\mathcal{H}_i$, the change in $S_I(r)$ can be expressed by

$$\frac{dS_I(r)}{dc_o} = \mu_{n-2}(\mathcal{F}_i) \cdot \frac{r}{r_1} \tag{III.26}$$

where

$$\mathcal{F}_i = \bigcap_{i=1}^{n+1} \mathcal{H}_i \cap \mathcal{S}_o(r) \cap \{X: X \cdot \hat{X}_i = c_o\}, \tag{III.27}$$

$r_1 = \left(r^2 - c_o^2\right)^{1/2}$, and μ_{n-2} is (n-2)-dimensional measure. In other words, each differential element of the intersection $\mathcal{F}_i$ sweeps out an increment of surface on a cylinder with base $\mathcal{F}_i$. The difference between the cylinder surface generated and the intersected shell content is the ratio r/r_1, which is the same for all differential elements. For unrestricted motion of $\mathcal{H}_i$, we have

$$dS_I = \int_{\mathcal{F}_i} \frac{r}{r_1} m(Y)\, dS_{n-1}(Y) \qquad \text{(III.28)}$$

where $m(Y)$ is the displacement of $\mathcal{H}_i$ at the point Y $(Y \in \mathcal{H}_i - c_i \cdot \hat{X}_i)$. The ratio r/r_1 is again constant over $\mathcal{F}_i$, and $m(Y) = c + L(Y)$ where L is a linear function because it corresponds to movement of a hyperplane.

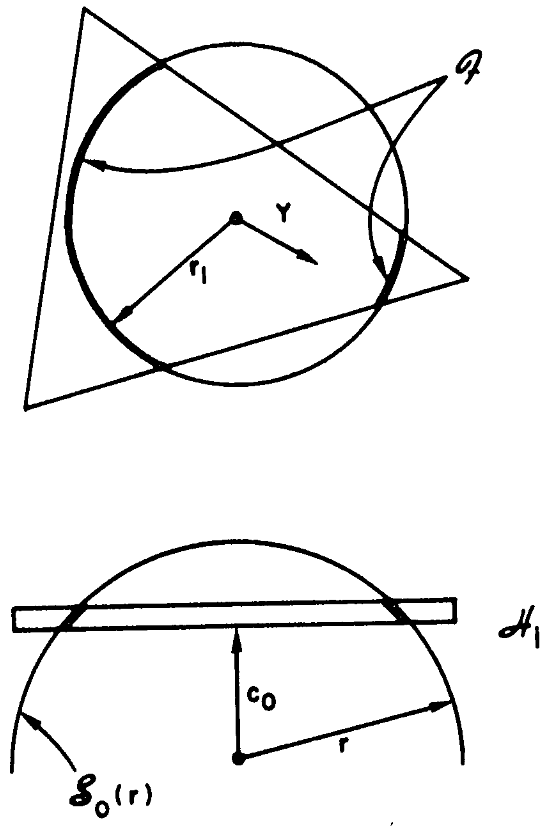

Fig. 24.

$$dS_I = \frac{r}{r_1} \int_{\mathcal{F}_i} (k_o + L(Y))dS_{n-1}(Y)$$

$$= \frac{r}{r_1}\left[k_o \int_{\mathcal{F}_i} dS_{n-1}(Y) + L\left(\int_{\mathcal{F}} YdS_{n-1}(Y)\right)\right]$$

$$= \frac{r}{r_1} \int_{\mathcal{F}_i} dS_{n-1}(Y)\left[k_o + L\left(\frac{\int_{\mathcal{F}} YdS_{n-1}(Y)}{\int_{\mathcal{F}} dS_{n-1}(Y)}\right)\right] \qquad \text{(III.29)}$$

Denote by $\overline{Y_i}$ the argument of L in this last expression. $\overline{Y_i}$ is just the center of mass in the face $\mathcal{H}_i$ of $\mathcal{F}_i$; whence $k_o + L(\overline{Y_i})$ is just the orthogonal displacement of the center of mass and the above can be interpreted as

$$\frac{r}{r_1} \cdot \text{(mass of } \mathcal{F}_i\text{) (displacement of center of mass)} \qquad \text{(III.30)}$$

Returning now to Conjecture 3, suppose that for one of the $\mathcal{H}_i$, $\overline{Y_i}$ is not 0; that is, the center of mass of $\mathcal{F}_i$ is not at the tip of $c_o\hat{X}_i$. Then we perform the mathematical analogue of putting our thumb on $\overline{Y_i}$ and letting $\mathcal{H}_i$ roll along the surface of the hypersphere $|X| \leq c_o$. More precisely, let

$$c_o\hat{X}_i'(\alpha) = c_o a(\alpha)\hat{X}_i + \alpha\overline{Y_i} \qquad \text{(III.31)}$$

where $a(\alpha)$ is chosen to keep $|\hat{X}_i'| = 1$.

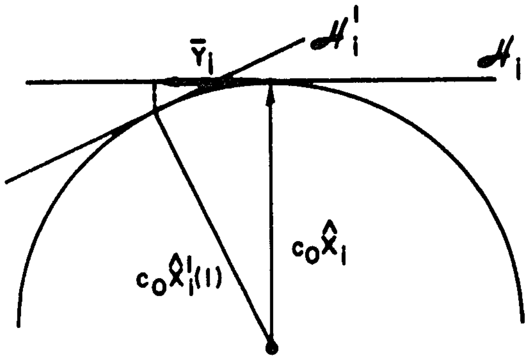

Fig. 25.

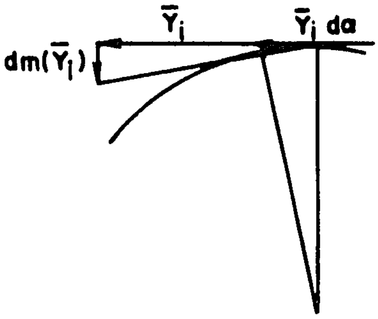

Fig. 26.

$$\left.\frac{dS_I}{d\alpha}\right|_{\alpha=0} = \beta \left.\frac{dm(\bar{Y}_i)}{d\alpha}\right|_{\alpha=0} = -\beta\frac{|\bar{Y}_i|^2}{c_o|X_i|} < 0 \qquad \text{(III.32)}$$

The constant β is greater than zero unless $\mathcal{F}_i = \emptyset$.

In sum, if the center of mass is not at $c_o\hat{X}_i$ then we can arrange to move it inward, cutting off more of the shell, while still satisfying the hyperplane distance constraint. This leads to:

Lemma 4. Assuming the vectors $\hat{X}_i$ are not all contained strictly within a halfspace (i.e., there does not exist a vector $\hat{X}_o$ such that $\hat{X}_o \cdot \hat{X}_i \geq 0$ for all $i = 1,\ldots,n+1$), a sufficient condition for Conjecture 3 is that, given $n-1$ vectors $Y_1,\ldots,Y_{n-1}$ in $(n-1)$-dimensional space and a shell $\mathcal{S}_o(r)$, there is at most one additional vector Y_n such that

$$\int_{\mathcal{D}} Y dS_{n-1} = \vec{0} \qquad \text{(III.33)}$$

where $\mathcal{D}$ is the intersection of $\mathcal{S}_o(r)$ with the convex null of the Y_i. In other words, there is at most a unique position for the remaining vertex such that the polytope $Y_1 \ldots Y_n$ has this intersected center of mass at the origin.

Proof. Define the matrix $\hat{\underline{X}}_i$ to be a row of all the column vectors $\hat{X}_j$, $j \neq i$. The assumption on the $\hat{X}_i$ implies that $\hat{\underline{X}}_i^T$ is non-singular for each i since if

$$\hat{\underline{X}}_i^T \cdot Z = \vec{0} ,$$

then either $Z/|Z|$ or $-Z/|Z|$ will serve as a violating $\hat{X}_o$. The rank of $\hat{\underline{X}}_i$ being equal to the rank of $\hat{\underline{X}}_i^T$, we conclude that the equation

$$\hat{\underline{X}}_i \cdot Z = \begin{bmatrix} c_o \\ \vdots \\ c_o \end{bmatrix} \tag{III.34}$$

has an unique solution, say Z_i. Note that by definition the Z_i are the extreme points of the polytope $\bigcap_i \mathcal{H}_i$. Furthermore, every set of n Z_i's must be independent. To see this, let $\underline{Z}_i$ be the row of column vectors Z_j, $j \neq i$. Then

$$Y \cdot \underline{Z}_1^T = 0 \implies Y \cdot \underline{Z}_1^T \cdot \underline{X}_j = 0 \tag{III.35}$$

Assuming

$$\underline{Z}_1 = [Z_2 \ldots . \quad \ldots . \quad \ldots . Z_{n+1}]$$

and

$$\underline{X}_j = [X_1 \ldots X_{j-1} X_{j+1} \ldots X_{n+1}]$$

$$\underline{Z}_1^T \cdot \underline{X}_j = (\beta_{im}) \qquad \beta_{im} = c_o \qquad i \neq m+1$$
$$\beta_{i,i-1} < 0 \tag{III.36}$$

The constant $\beta_{i,i-1}$ is $Z_i \cdot X_i$, which must be negative or else $Z_i/|Z_i|$ violates the constraint.

Letting $Y = (y_1, y_2, \ldots, y_n)$, Eq. (III.35) can thus be written

$$y_i \beta_{i,i-1} + c_o \sum_{m \neq i} y_m = 0, \quad i = 2,\ldots,n, \quad c_o \sum_{i=1} y_i = 0 .$$

$$\Rightarrow \text{ for } i = 2,\ldots,n \quad \sum_{m \neq i} y_m = -y_i \Rightarrow y_i \beta_{i,i-1} = c_o y_i$$

$$\Rightarrow y_i = 0, \quad i = 1,2,\ldots,n, \quad \text{since} \quad c_o > 0 . \qquad \text{(III.37)}$$

To return to the main stream of the proof, it is well known that the regular simplex minimizes $\overline{R}/\underline{R}$, where $\overline{R}$ is the radius of the smallest hypersphere containing the simplex, and $\underline{R}$ that of the largest hypersphere contained within. (See, for example, [16] or the next lemma.) We may therefore restrict our attention to the non-trivial cases where at least one of the Z_i has $|Z_i| > r$ and, of course, $c_o < r$. This, plus the independence of the Z_i, ensures that for at least one of the $\mathcal{H}_i$, $\mathcal{F}_i \neq \emptyset$. Although $\mathcal{F}_i$ was originally defined in terms of the $\mathcal{H}_i$, it is also expressible as the intersection of a (n-1)-dimensional shell $\mathcal{S}_o(r)$ with the convex hull of the Z_j's in $\underline{Z}_i$. By the remarks preceding this lemma, the center of mass of $\mathcal{F}_i$ must be at the origin of the subspace $\mathcal{H}_i - c_o \hat{X}_i$. Letting Y be a generic vector of the subspace, the condition is

$$\int_{\mathcal{F}_i} Y dS_{n-1} = \vec{0}. \qquad \text{(III.38)}$$

Now, suppose all the Z_j in $\underline{Z}_i$ are fixed except for Z_ℓ. The hypothesis of the lemma implies that there is at most a unique Z_ℓ such that (III.38) is satisfied. But if we then examine $\mathcal{F}_\ell$, we observe that all the Z_j, $j \neq \ell$, $j \neq i$, are common to both $\mathcal{F}_i$ and $\mathcal{F}_\ell$, and moreover, since $Z_j \cdot c_o\hat{X}_i = Z_j \cdot c_o\hat{X}_i$ for these vectors, they are in the same relation to the origin of $\mathcal{H}_\ell - c_o\hat{X}_\ell$ as to the origin $\mathcal{H}_i - c_o\hat{X}_i$. Since the shells forming the $\mathcal{F}_j$'s are all the same size, applying condition (III.38) and the uniqueness requirement leads us to conclude that $|Z_\ell - Z_j| = |Z_i - Z_j|$ for all $j = 1,\dots,n+1$. The indices i and ℓ being arbitrary, we have that all two-dimensional edges are the same length, and therefore the Z_i's form a regular simplex. ∎

Comprehension of the n-dimensional proof will undoubtedly be facilitated by a three-dimensional example. Consider a sphere intersected by four planes each at distance c_o from the origin (Fig. 27). Examining the face of $Z_1Z_2Z_3$ we see a circle intersecting a triangle (Fig. 28). The center of mass condition says that the center of mass of the interior perimeter must be at the center of the circle, and the uniqueness condition says that given Z_1 and Z_2, Z_3 can be in only one position for this to occur. Now we flip around the corner and examine $Z_1Z_3Z_4$. Because the

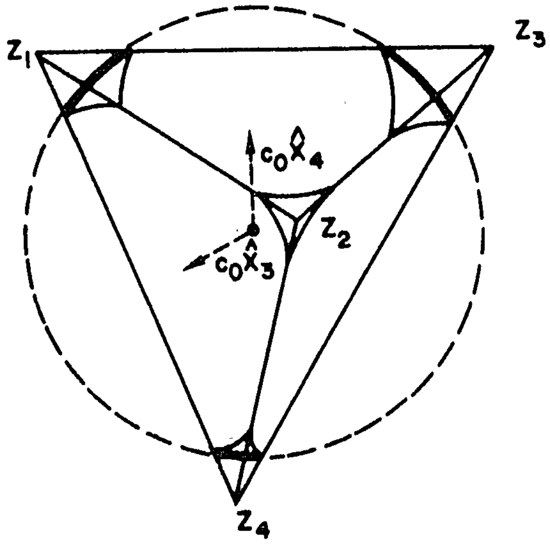

Fig. 27.

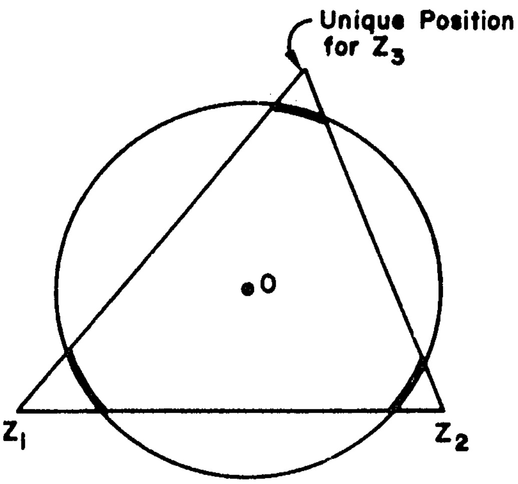

Fig. 28.

planes are at the same distances, the size of the circle is the same. Furthermore

$$Z_i \cdot \hat{X}_4 = Z_i \cdot \hat{X}_3 = c_o$$

$$\Rightarrow \quad |Z_i - c_o\hat{X}_4| = |Z_i - c_o\hat{X}_3| \qquad \text{for} \quad i = 1,2. \qquad \text{(III.39)}$$

As a result, the uniqueness condition provides that the center of mass in Z_1, Z_2, Z_4 will be at $c_o\hat{X}_3$ only if Z_4 is located where Z_3 would land if we took $Z_1Z_2Z_3$ and swung it about the line Z_1Z_2 until it coincided with the plane of $Z_1Z_2Z_4$. Consequently, we must have

$$|Z_2 - Z_3| = |Z_2 - Z_4| \qquad \text{and} \qquad |Z_1 - Z_3| = |Z_1 - Z_4| .$$

We should comment that the halfspace restriction on the $\hat{X}_i$ is not unduly restrictive. Balakrishnan [8] has shown that in the Gaussian case only signal sets whose convex hull contains the origin can be optimal. His argument is readily adapted to the general situation for it demonstrates a strict dominance. Hence any possible optimal signal set can be approximated arbitrarily closely by a signal set satisfying the assumption of Lemma I. Continuity of the probability of error as a function of the signal set then guarantees that if the regular simplex is the only solution indicated by the lemma, its performance must be within δ of that of any candidate for optimality, with $\delta > 0$ but arbitrarily small. The restriction undermines the uniqueness but not the global optimality of the simplex solution.

For all of its aesthetic appeal, Lemma I seems to offer little toward the solution of the simplex conjecture. Whereas we have traded in an n-dimensional problem for an (n-1)-dimensional problem, it is questionable whether the complexity of the new formulation, understanding that it must be verified for all dimensions, is any less. Nevertheless, it will furnish, with a modicum of additional effort, yet another proof of the conjecture in three dimensions.

As a means of comparison and a consolation, we present the following known result which projects a thin ray of light into the n-dimensional gloom.

Lemma 5. Of all simplexes of content V_o which intersect a given hypersphere, the centered regular simplex maximizes the intersected content.

Proof:[1] Clearly the intersection of one line segment with another is maximized by centering one on top of the other. Given an n-dimensional simplex, slice it up along hyperplanes parallel to one face of the simplex (Fig. 29). Applying the induction hypothesis in each slice, we maximize the intersection in each slice uniformly. Linearity of the sections of a simplex permits us to reassemble the slices into a simplex. Moreover, unless the simplex was originally symmetric in this sense, we will have strictly increased the intersection content. Convergence to the regular simplex is guaranteed, because this content is overbounded by V_o. ∎

The spherical analogue of this lemma, that of all spherical simplexes of solid angle Ω, the centered regular spherical simplex maximizes the intersection with any cap is almost sufficient to prove Conjecture 3. The added ingredient is simply concavity of intersection as a function of Ω for any fixed cap.

[1]This proof was mentioned to us by Dr. T. M. Cover.

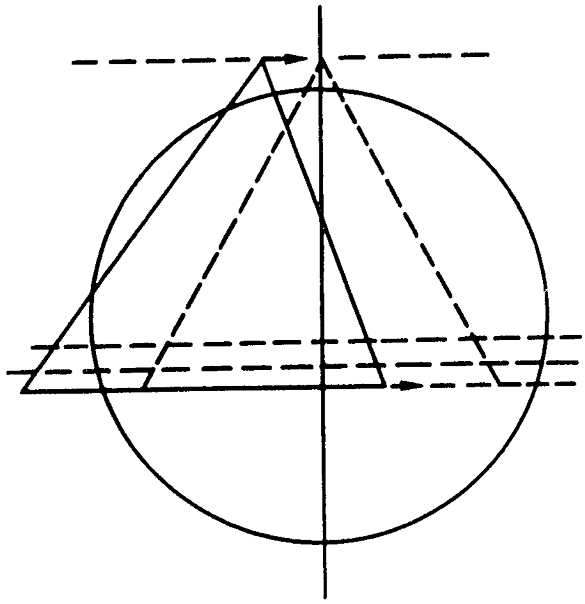

Fig. 29.

At long last we are in a position to expound the relative merits of the foregoing conjectures. First, we note that Conjecture 1 easily implies Conjecture 2:

$$\int_0^{r'} t(r)dr = \int_0^{r'} \tilde{t}(r)dr \Rightarrow r' \geq r_o \quad \text{of Conjecture 1}$$
$$\Rightarrow \tilde{t}(r') \leq t(r') \tag{III.40}$$

Conjecture 2 in (n-1)-dimensions combined with Lemma 1 of Chapter II and a straightforward convexity argument comprises a proof of the third conjecture in n-dimensions. However, it is also true that Conjecture 2 in n-dimensions immediately gives Conjecture 3 in n-dimensions: Balakrishnan [9] has shown that the regular simplex maximizes the minimum distance between signals. This enables us to say that Conjecture 3 is valid for c_o sufficiently large. In other words, at some c_o the caps $X \cdot \hat{X}_i \geq c_o$ $i = 1,\ldots,n+1$, $|X| = r_o$, will be disjoint for the regular simplex, whereas they are not for any other configuration. At this point, $\tilde{t}(r_o) < t(r_o)$, implying the condition of Conjecture 2,

$$\int_0^{r_0} t(r)dr < \int_0^{r_0} \tilde{t}(r)dr \;, \qquad \text{(III.41)}$$

and therefore $t(r) \le t(r)$ for $r > r_0$ as well as for $r \le r_0$.

The moral of the story is that in choosing to attack the problem via one of the preceding conjectures, the only significant criterion is the adaptability of the characterization to an induction proof. In terms of establishing a proof for arbitrary dimension, the effort of Lemma 1 of Chapter II, and its counterparts in the works of Landau and Slepian and Farber, is of little use unless it plays a role in the induction.

Chapter IV

A LINEARIZED VERSION

Taking the ball density conjecture (Eq. (I.13)) as a point of departure, we propose in this chapter a class of problems of which the ball density conjecture, and thus the simplex code conjecture, is a special case. Among these problems we find a "linearized" version of the ball density conjecture for which we provide a proof.

Let $\mathcal{B}_i$ be the set $\{X: |X-\hat{S}_i| < r_o\}$ for each of the $n+1$ $\hat{S}_i$'s. If $I_i(X)$ is the indicator function of $\mathcal{B}_i$, we define intersection sets $\mathcal{A}_j$ by

$$\mathcal{A}_j = \left\{X: \sum_{1}^{n+1} I_i(X)=j\right\} \tag{IV.1}$$

so that $\mathcal{A}_j$ is the set of points contained in exactly j of the $\mathcal{B}_i$. Letting A_j be the measure of the set $\mathcal{A}_j$, form the vector

$$A = (A_1, \ldots, A_{n+1}) \tag{IV.2}$$

We have in this way constructed a function

$$f: \mathbb{E}^{n\cdot(n+1)} \to \mathbb{E}^{n+1} \tag{IV.3}$$

which maps a set of vectors $\hat{S}_1, \ldots, \hat{S}_{n+1}$ into a vector A. It is defined on the $(n+1)$-fold Cartesian product of the closed n-dimensional unit shell and can readily be seen to be continuous. Consequently the range of $f(\underline{S})$ as the matrix $\underline{S} = (\hat{S}_1, \ldots, \hat{S}_{n+1})$ varies over all allowable signal sets is a compact set in $\mathbb{E}^{n+1}$, say $\mathcal{A}$.

One can then ask any number of questions regarding the properties of this set $\mathcal{A}$. First we note that the definition of $\mathcal{A}$ requires that

$$A \cdot (1,2,\ldots, n+1) = A_1 + 2A_2 + \ldots + (n+1)\, A_{n+1} \qquad \text{(IV.4)}$$

be a constant for all $A \in \mathcal{A}$, since this is merely the sum of contents of the $n+1$ balls. Now we can inquire for what linear functionals, $Z \in \mathbb{E}^{n+1}$, does the regular simplex achieve the extremum for the set $\mathcal{A}$. In this context the ball density conjecture is that

$$A \cdot (1,\ldots, 1), \quad A \in \mathcal{A} \qquad \text{(IV.5)}$$

is maximized by A^*, the vector corresponding to the regular simplex.

It is clear that our only interest in this way of viewing the problem is to find a technique for proving the conjecture indirectly by exposing some more salient feature of the geometry involved. To date all attempts have been frustrated. However, we can prove that

$$A \cdot \left(0, \binom{2}{2}, \ldots, \binom{n+1}{2}\right) \qquad \text{(IV.6)}$$

is minimized by A^*. This functional has been deliberately chosen to represent a linearized version of the ball density conjecture, as we will now show. Consider two n-dimensional balls of equal content B_o and centers X_1 and X_2.

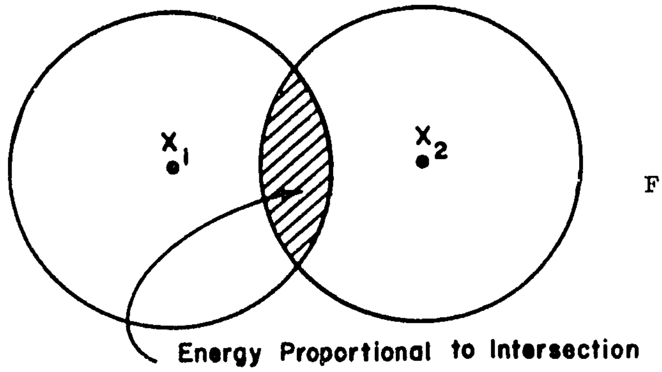

Fig. 30

Suppose there exists a repelling force between the centers such that the energy required to bring the second ball from infinity to X_2 is equal to the content of the intersection $\mathcal{B}(X_1) \cap \mathcal{B}(X_2)$. If the balls have radius r_o, it is easy to see that the force is zero if $|X_1 - X_2| \geq 2r_o$ and increases monotonically to B_o as $|X_1 - X_2|$ tends to zero. Now suppose such a repelling force exists between each of the signal vectors $\hat{S}_i$. It is not difficult to verify that the energy of any signal set is just the expression (IV.6).

The qualitative difference between this problem and the probability of error problem is that for the former the effect of one ball on another is not altered by the presence of a third ball, whereas for the latter it is. Thus in this sense it is the linearized version of the conjecture.

We now prove that if the signals are acted upon by any force which is non-increasing in the separating distance, the regular simplex is the minimum energy configuration.

Theorem 2. Let the force repelling two points, X_1 and X_2, be $f(\ell)$, where $\ell = |X_1 - X_2|$ and $f(\ell)$ is a non-increasing function of ℓ such that for any $\epsilon > 0$,

$$\int_{\epsilon}^{\infty} f(\ell)\, d\ell < \infty \,. \tag{IV.7}$$

Then the energy, E, required to bring $n+1$ points from infinity to positions $\hat{S}_1, \ldots, \hat{S}_{n+1}$ is a minimum if the $\hat{S}_i$ are the vertices of a regular simplex.

Proof. Let

$$\ell_{ij} = |\hat{S}_i - \hat{S}_j| \,. \tag{IV.8}$$

By linearity we can write

$$E = \sum_{j>i} e(\ell_{ij}) \tag{IV.9}$$

where
$$e(t) = \int_t^\infty f(\ell)\ d\ell \tag{IV.10}$$

The assumption on f implies that $e(t)$ is convex, so that

$$E \geq \frac{(n+1)\ n}{2}\ e(\bar{\ell}) \tag{IV.11}$$

where $\bar{\ell}$ is the average distance between points,

$$\bar{\ell} = \frac{2}{(n+1)\ n} \sum_{j>i} \ell_{ij} \tag{IV.12}$$

The distance between $\hat{S}_i$ and $\hat{S}_j$ can be expressed

$$\ell_{ij} = \ell(\alpha_{ij}) \tag{IV.13}$$

where
$$\ell(\alpha) = \sqrt{2-2\alpha} \tag{IV.14}$$

and
$$\alpha_{ij} \equiv \hat{S}_i \cdot \hat{S}_j \tag{IV.15}$$

Recalling the definition of $\underline{S}$ as the row of column vectors $\hat{S}_i$, it is easy to see that $\underline{\alpha}$ is non-negative definite.

$$\alpha \equiv \underline{S}^T\underline{S} \Rightarrow X^T\underline{S}^T\underline{S}X = (\underline{S}X)^T\ \underline{S}X \geq 0\ . \tag{IV.16}$$

It is symmetric and all diagonal elements are one. Consequently, taking X to be the all ones vector in the above yields

$$\sum_{i>j} \alpha_{ij} \geq \frac{-\ (n+1)}{2} \tag{IV.17}$$

By the concavity of $\ell(\cdot)$

$$\bar{\ell} = \frac{2}{(n+1)\ n} \sum_{j>i} \ell_{ij} = \frac{2}{(n+1)\ n} \sum_{j>i} \ell(\alpha_{ij}) \leq \ell(\bar{\alpha}) \qquad \text{(IV.18)}$$

where $\bar{\alpha} = \frac{2}{(n+1)\ n} \sum_{i>j} \alpha_{ij}$ and equality holds if and only if all the $\alpha_{ij} = \bar{\alpha}$. But (IV.17) implies that

$$\bar{\alpha} \geq -\frac{1}{n}, \qquad \text{(IV.19)}$$

and since decreasing $\bar{\alpha}$ increases $\ell(\bar{\alpha})$, we have

$$\bar{\ell} \leq \ell\left(-\frac{1}{n}\right) \qquad \text{(IV.20)}$$

with equality if and only if the signals form a regular simplex. Returning to the energy function, we have

$$E \geq \frac{(n+1)\ n}{2}\, e(\bar{\ell}) \geq \frac{(n+1)\ n}{2}\, e\left(\ell\left(-\frac{1}{n}\right)\right). \qquad \text{(IV.21)}$$

This last expression is precisely the energy of the regular simplex configuration. Either the first or the second inequality is strict unless trivially $E = 0$ or the signals form a regular simplex. ∎

The restriction on the force law is necessary to preclude anomalies such as the following:

Suppose three particles constrained to lie on the unit circle are repelled by a force

$$\begin{aligned} f(\ell) &= 0 \qquad \ell \leq 2\sqrt{3}, \quad \ell > 2 \\ &= 1 \qquad 2\sqrt{3} < \ell \leq 2 \end{aligned}$$

Then the minimum energy configuration is two particles at one end of a diameter and one at the other.

Admittedly our theorem finds natural application in the field of n-dimensional electrostatics rather than communication theory. It does, however, extend the maximum minimum distance idea, as well as making an interesting comparison with the simplex conjecture itself.

Chapter V

GEOMETRICAL INTERPRETATION OF THE RELATION BETWEEN MEAN WIDTH AND OPTIMALITY FOR LOW SIGNAL-TO-NOISE RATIOS

In this chapter we assume the noise density to be fixed and consider the transmitted signal corresponding to the i^{th} message to be $\lambda \hat{S}_i$. Balakrishnan [7] has shown that for white Gaussian noise the derivative of the probability of correct decoding with respect to λ is

$$\frac{d\,\Pr(\text{correct};\lambda)}{d\lambda} = \frac{B}{(2\pi)^{\frac{1}{2}(n-1)}} \int_0^\infty r^{n-1} \exp\left(-\frac{r^2}{2}\right) dr \qquad (V.1)$$

where

$$B = \int_\Omega H_S(\hat{X})\, d\Omega = \int_\Omega d\Omega \cdot \overline{B} \qquad (V.2)$$

and $H_S(\hat{X})$ is the support function of the polytope formed by the signal vectors;

$$H_S(\hat{X}) = \max_i \hat{X} \cdot \hat{S}_i \qquad (V.3)$$

Hence maximizing the mean width $\overline{B}$ of this polytope maximizes the above derivative, which in turn implies global optimality as $\lambda \to 0$. Although the derivation is for Gaussian noise statistics, a similar statement could be made for general monotonically decreasing spherically symmetrical noise functions, $f(|Z|)$, by expanding $f(|Z|)$ in a series about 0. We would now like to show the geometrical reason for this relationship.

First we rewrite the integral B as

$$B = \sum_{1}^{n+1} \int_{\mathcal{R}_i} \hat{X} \cdot \hat{S}_i \, d\Omega \tag{V.4}$$

where $\mathcal{R}_i$ is the region where $\hat{X} \cdot \hat{S}_i$ is the maximum. The contribution of each $\hat{S}_i$ is

$$\int_{\mathcal{R}_i} \hat{X} \cdot \hat{S}_i \, d\Omega = \int_{\mathcal{R}_i} \cos\theta \, d\Omega \quad . \tag{V.5}$$

$$\theta \equiv \cos^{-1} \hat{X} \cdot \hat{S}_i \tag{V.6}$$

Now consider a ball noise density

$$\begin{aligned} f(|Z|) &= K \qquad |Z| < 1 \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{V.7}$$

The probability of correct decoding can be written

$$P_c(\lambda) = \frac{1}{n+1} \int f(\min_i |X - \lambda \hat{S}_i|) \, dX$$

$$= \frac{1}{n+1} \sum_i \int_{\mathcal{R}_i} f(|X - \lambda \hat{S}_i|) \, dX \tag{V.8}$$

since the minimum can be seen to coincide with the maximum of $X \cdot \hat{S}_i$. Examining the i^{th} region, the integral can also be written

$$\int_{\mathcal{R}_i} f(|X - \lambda \hat{S}_i|) \, dX = K \int_{\mathcal{R}_i} (\ell(\theta))^n \, d\Omega \tag{V.9}$$

where we assume $\lambda < 1$.

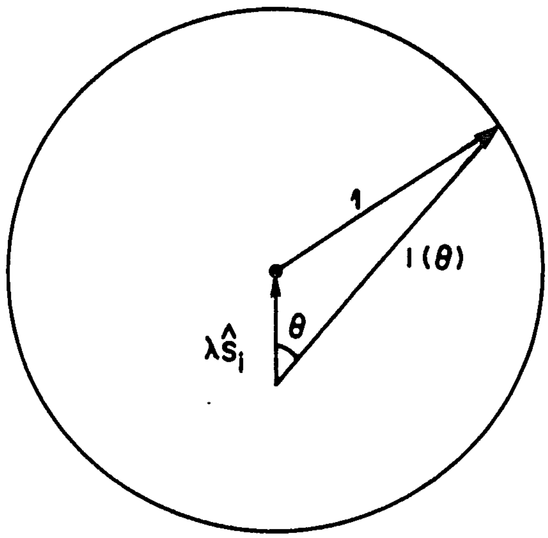

Fig. 31

In other words, if X makes an angle θ with $\hat{S}_i$, then the probability in a ray of size $d\Omega$ along Y is proportional to the volume of the intersection of the ray with the ball. This volume is just $C(\ell(\theta))^n \, d\Omega$, where $\ell(\theta)$ is the distance from the origin to the surface of the ball at the angle θ. But now, recalling $|\hat{S}_i| = 1$,

$$1 + \lambda \cos \theta - \lambda^2 \sin^2 \theta \leq \ell(\theta) \leq 1 + \lambda \cos \theta \qquad \text{(V.10)}$$

Thus $$\frac{d}{d\lambda} \int_{\mathfrak{R}_i} (\ell(\theta))^n \, d\Omega = \int_{\mathfrak{R}_i} \frac{d}{d\lambda} (1 + n\lambda \cos \theta + \text{higher order}) \, d\Omega \qquad \text{(V.11)}$$

Evaluating at $\lambda = 0$,

$$= n \int_{\mathfrak{R}_i} \cos \theta \, d\Omega \qquad \text{(V.12)}$$

Hence as $\lambda \to 0$ a unit ball density assigns the same weight to every differential element $d\Omega$ as does the support function of the mean width integral. Since any general $f(|Z|)$ can be uniformly approximated by ball densities, the mean width and global optimality as $\lambda \to 0$ are thus intimately related.

We note, however, that although several authors, [7], [8], [10], have assumed the existence of a proof that the regular simplex maximizes the mean width over all polytopes with $n+1$ vertices on the unit hypersphere, we have been unable to locate this proof. Indeed, the only theorem we have discovered is that in three dimensions, the regular simplex <u>minimizes</u> the mean width over all simplexes of equal content.[1]

---

1. Personal correspondence from Professor H. Busemann, Department of Mathematics, University of Southern California.

Chapter VI

CONCLUSIONS

While the quixotic objective of our research has always been, of course, the proof of the simplex conjecture as given in Eq. (I.4), our efforts have served primarily to define more clearly the essence of its complexity, and in doing so, to suggest new directions for future work. The reexamination reveals that the Landau and Slepian inequality rests heavily on the single hyperplane requirement. We regret to say that in spite of its ostensible intricacy, the theorem yields an inequality which is far too weak. Single hyperplanes intersecting hyperspheres represent a fairly simple geometry. To solve the simplex conjecture along these lines requires an intimate understanding of how arbitrary simplexes intersect hyperspheres. The theorem's best use seems to be in refining the lower bound for the probability of error when the block length M is finite. The basic sphere-packing argument ([4], pp. 618-621) assumes the solid angle allotted to a particular signal's decoding region to be concentrated in a cap about the signal. A refinement could be achieved with the inequality by using a cap less n single hyperplane regions rather than just a cap. In this application, the inequality can be viewed as a first order correction to the overlap error.

We have endeavored to develop the detailed knowledge of the nature of the intersection of a simplex and a hypersphere by means of three conjectures. The first two differ substantially from those of previous authors, and both can be used to prove the extended Fejes-Toth theorem. For the third we prove a markedly new sufficient condition. We had hoped that a proof for at least the fourth dimension would materialize, as all three approaches reduce it to a three-dimensional problem, but our attempts have been rebuffed by a series of special cases. In terms of the n-dimensional problem, it appears that all formulations lead to the same impasse, demonstration of a unique solution to n simultaneous nonlinear equations in n unknowns. The n-dimensional geometry seems rich enough to assure the non-degeneracy of these equations and prevent a successful induction proof.

The importance of the nonlinearity is underlined by our proof of a linearized version of the conjecture. This result can be seen to include the maximum minimum distance characterization of Balakrishnan ([9], pp. 30-31). To forestall possible alarm, we should mention that the minimum distance proofs of both [9] and [10] are unnecessarily long, since for any discrete random variable α

$$E(\alpha) = -\frac{1}{M-1} \quad \text{and} \quad \max \alpha = -\frac{1}{M-1}$$

$$\Longrightarrow \qquad \alpha \equiv -\frac{1}{M-1} \quad .$$

After showing a geometrical relation between the mean width of the polytope with the signal set as vertices and optimality as $\lambda \to 0$, we comment that the maximum mean width problem may still be open. The known result, minimum mean width over all simplexes of equal content for $n = 3$, may be extendable to higher dimensions, but it appears to be far weaker than the required condition.

Although we failed to resolve the conjecture, it is hoped that this thesis will make an appreciable contribution to the search for a proof by serving as a sounding board for new ideas. It is easy to mistake a new formulation for real progress. By providing a detailed basis of comparison, our work may facilitate recognition of true novelty as well as suggesting some paths previously unexplored.

APPENDIX A

This appendix consists of a verbatim statement and proof of a lemma proved in Appendix A of [5], pp. 1261-1262. The equation numbers have been changed.

Lemma: Let $w_1(x)$ and $w_2(x)$ be integrable functions that satisfy

$$\int_a^b w_1(x)dx = \int_a^b w_2(x)dx. \tag{A.1}$$

Further, suppose there exists an x', $a \leqq x' \leqq b$, such that

$$\begin{aligned} w_2(x) &\geqq w_1(x), \qquad a \leqq x \leqq x' \\ w_2(x) &\leqq w_1(x), \qquad x' \leqq x \leqq b. \end{aligned} \tag{A.2}$$

Then, if $m(x)$ is a nonnegative monotone increasing function,

$$\int_a^b m(x)w_1(x)dx \geqq \int_a^b m(x)w_2(x)dx. \tag{A.3}$$

If $m(x)$ is a nonnegative monotone decreasing function,

$$\int_a^b m(x)w_1(x)dx \leqq \int_a^b m(x)w_2(x)dx. \tag{A.4}$$

Equality holds in (A.3) and (A.4) only if $w_1(x) = w_2(x)$ for almost all x.

Proof: If $m(x)$ is nonnegative and monotone increasing, then

$$\int_a^b m(x)[w_1(x) - w_2(x)]dx$$

$$= \int_a^{x'} m(x)[w_1(x) - w_2(x)]dx + \int_{x'}^b m(x)[w_1(x) - w_2(x)]dx$$

$$\geq m(x') \int_a^{x'} [w_1(x) - w_2(x)]dx + m(x') \int_{x'}^b [w_1(x) - w_2(x)]dx$$

$$= m(x') \left[\int_a^b w_1(x)dx - \int_a^b w_2(x)dx \right] = 0.$$

If $m(x)$ is nonnegative and monotone decreasing, the steps are the same with the inequalities reversed.

REFERENCES

1. Shannon, C. E., "Communication in the Presence of Noise," Proc. IRE, 37, Jan 1949, pp. 10-21.

2. Kotel'nikov, V. A., thesis, Molotov Energy Institute, Moscow, 1947, translated by R. A. Silverman as The Theory of Optimum Noise Immunity, McGraw-Hill Book Co., New York, 1959.

3. Wozencraft, J. M., and Jacobs, I. M., Principles of Communication Engineering, John Wiley & Sons, New York, 1967.

4. Shannon, C. E., "Probability of Error for Optimal Codes in a Gaussian Channel," BSTJ, 38, May 1959, pp. 611-656.

5. Landau, H. J., and Slepian, D., "On the Optimality of the Regular Simplex Code," BSTJ, 45, 8, Oct 1966, pp. 1247-1271.

6. Sommerville, D. M. Y., An Introduction to the Geometry of N Dimensions, Dover Publications, Inc., New York, 1958, pp. 96-97.

7. Gilbert, E. N., "A Comparison of Signaling Alphabets," BSTJ, 31, May 1952, pp. 504-522.

8. Balakrishnan, A. V., "A Contribution to the Sphere-Packing Problem of Communication Theory," Jour. of Math. Anal. and Appl., 3, 3, Dec 1961, pp. 485-506.

9. Balakrishnan, A. V., "Signal Selection Theory for Space Communication Channels," Chapter 1 in Advances in Communication Systems, A. V. Balakrishnan, editor, Academic Press, New York, 1965.

10. Weber, C. L., Elements of Detection and Signal Design, McGraw-Hill Book Co., New York, 1968.

11. Wyner, A. D., "On the Probability of Error for Communication in White Gaussian Noise," IEEE Trans. on Inf. Theory, IT-13, 1, Jan 1967, pp. 86-90.

12. Fejes-Toth, L., Lagerungen in der Ebene auf der Kugel und im Raum, Springer-Verlag, Berlin, 1953, pp. 137-141.

13. Farber, S. M., "On the signal selection problem for phase coherent and incoherent communication channels," Communications Theory Lab., California Institute of Technology, Pasadena, Calif., Tech. Rept. 4, May 1968.

14. Schaffner, C. A., and Kreiger, H. A., "The Global Optimization of Two and Three Phase-Incoherent Signals," Technical Rept. No. 3, Communications Theory Lab., California Institute of Technology, Jan 1968.

15. Blachman, N. M., "Geometry of Optimum Incoherent Detection," IEEE Trans. on Inf. Theory, IT-16, 2, Mar 1970, pp. 202-205.

16. Slepian, D., "The Content of Some Extreme Simplexes," Pac. Jour. of Math., 31, 3, 1969, pp. 795-808.