QUARTERLY OF APPLIED MATHEMATICS
APRIL 1984



# A SIMPLICIAL MAPPING METHOD FOR LOCATING THE ZEROS OF A FUNCTION*

BY

C. S. HSU (*University of California, Berkeley*)

AND

W. H. ZHU (*Zhejiang University, People's Republic of China*)

**Abstract.** The currently available methods of finding the zeros of a vector function are quite effective but they do require a prior knowledge of the approximate locations of the zeros. Such a knowledge is, however, often not available, hence making the task of finding the zeros very difficult. Developed in this paper is a method which enables us to search the whole domain of interest in a systematic way to locate *all* the simple zeros approximately. This method complements naturally the various conventional iteration procedures. The method involves (i) partitioning the domain into a system of simplexes and (ii) at each simplex examining a local simplicial mapping induced by the given function in order to test whether a zero is present. Special properties of the barycentric coordinates are used in devising a simple testing procedure which makes the method viable and effective.

**1. Introduction.** In analyzing physical systems mathematically, one is sometimes required to find all the zeros of a vector function in a certain domain of interest. For this purpose many methods and programs are available at the present time, including the well-known Newton's method and its variations [2, 4]. All these methods are capable of determining the zeros accurately by iteration if some information is available about their approximate locations. However, for most problems the approximate locations of the zeros are not always known and the determination of the zeros becomes often a hit and miss proposition.

In this paper we present a method which enables us to search throughout the domain of interest in a systematic and yet simple manner to locate approximately all the simple zeros which are apart by a certain distance. Once the approximate location of a zero has been determined by this search technique, any conventional iteration methods can then be used to improve the accuracy.

Let us consider an $n$-dimensional vector function $\mathbf{f}(\mathbf{x})$ defined over a euclidean $n$-space $X^n$. The present method is based upon the following idea. We first cover $X^n$ by a system of $n$-simplexes [1, 8]. At each $n$-simplex we approximate $\mathbf{f}(\mathbf{x})$ locally by a linear function $\mathbf{f}^L(\mathbf{x})$. This linear function is then examined to see whether its zero is in the simplex. Based

* Received October 15, 1982.

upon this information we infer whether $\mathbf{f}(\mathbf{x})$ has a zero in the neighborhood of this simplex. In this part of the analysis the barycentric coordinates [1, 8] play a key role. It turns out that the location of the zero of $\mathbf{f}^L(\mathbf{x})$ relative to the local simplex is characterized by the barycentric coordinates in a remarkably informative way. This characterization leads us then to a very simple testing procedure.

In Secs. 2–4 we first provide the necessary geometical background information on the properties of simplexes, barycentric coordinates, and the important matter of how to partition a given domain of interest into a system of $n$-simplexes. In Sec. 5 we present the main part of the analysis by discussing how the local linear approximating function is created and how to determine the location of its zero relative to the simplex itself. In the event that the linear approximating function is degenerate, one would be interested in knowing the location of the affine kernel subspace of the function relative to the simplex. This is discussed in Sec. 6. Finally, in Secs. 7 and 8 the operational aspects of the search method are discussed.

In the paper various basic aspects of the method are discussed in considerable detail, but the paper is not one on the algorithm and implementing program. These will be reported separately elsewhere.

**2. An $n$-simplex in $R^n$.** Consider a eulcidean $n$-space $R^n$ with a coordinate system $x_i$, $i = 1, 2, \ldots, n$. Let $s^n$ be a $n$-simplex in $R^n$ with vertices at $\mathbf{x}_j, j = 0, 1, \ldots, n$. For briefness of notation we shall use $\mathbf{x}_j$ to denote both the $j$th vertex point and the position vector of that point in $R^n$. In matrix representation $\mathbf{x}_j$ will always be taken to be a column vector; hence,

$$\mathbf{x}_j = \left[x_{1j}, x_{2j}, \ldots, x_{nj}\right]^T \tag{2.1}$$

where the superscript $[\cdot]^T$ denotes the transpose.

This set of vertices wil be called the vertex set $V$ of $s^n$. If $V_1 = \{\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_r}\}$ is a subset of $V$ of $r$ members, then we denote the complement set of $V_1$ in $V$ by $V_1^c$, or explicitly as $\{\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_r}\}^c$, which has $n - r + 1$ members. Sometimes it is also convenient to have a notation for sets of indices. Let $I(n)$ denote a set of indices $0, 1, \ldots, n$. If $I_1 = \{j_1, j_2, \ldots, j_r\}$ is a subset of $I(n)$ of $r$ members, then either $I_1^c$ or $\{j_1, j_2, \ldots, j_r\}^c$ will denote the complement set of $I_1$ in $I(n)$.

Any subset of $r + 1$ members of $V$ defines a $r$-simplex $s^r$. When it is desirable to exhibit the vertices of this simplex, we use the symbol $\Delta$ to indicate the ordered vertices involved,

$$s^r = \Delta(\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_r}). \tag{2.2}$$

To be definite, we shall always assume that $s^r$ of (2.2) is so defined that the vertices are ordered according to

$$j_0 < j_1 < \cdots < j_r. \tag{2.3}$$

A simplex with the same vertex set as $s^r$ but with a different vertex ordering will be a simplex having an equal or opposite orientation to $s^r$, as the case may be. With this convention our $s^n$ is defined as

$$s^n = \Delta(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n). \tag{2.4}$$

Associated with this simplex we define two basic matrices $\mathbf{\Phi}(\mathbf{x})$ and $\mathbf{\Phi}^+(\mathbf{x})$:

$$\mathbf{\Phi}(\mathbf{x}) = [\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n], \tag{2.5}$$

$$\mathbf{\Phi}^+(\mathbf{x}) = \begin{bmatrix} 1, 1, \ldots, 1 \\ \mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n \end{bmatrix}. \tag{2.6}$$

$\mathbf{\Phi}(\mathbf{x})$ is an $n \times (n+1)$ matric with the first, the second,..., and the $(n+1)$th columns occupied, respectively, by $\mathbf{x}_0, \mathbf{x}_1, \ldots,$ and $\mathbf{x}_n$, while $\mathbf{\Phi}^+(\mathbf{x})$ is the matrix $\mathbf{\Phi}(\mathbf{x})$ augmented at the top by a row of unit elements and, therefore, is an $(n+1) \times (n+1)$ matrix. When the $j$th column is deleted from $\mathbf{\Phi}(\mathbf{x})$, the resulting $n \times n$ matrix is denoted by $\mathbf{\Phi}_{(j)}(\mathbf{x})$

$$\mathbf{\Phi}_{(j)}(\mathbf{x}) = [\mathbf{x}_0, \mathbf{x}_1, \ldots, \hat{\mathbf{x}}_j, \ldots, \mathbf{x}_n] \tag{2.7}$$

where an overhead symbol $\hat{}$ denotes deletion. We shall also adopt the following notation for the determinants of $\mathbf{\Phi}^+(\mathbf{x})$ and $\mathbf{\Phi}_{(j)}(\mathbf{x})$

$$d(s^n) = \det \mathbf{\Phi}^+(\mathbf{x}), \qquad d_{(j)}(s^n) = (-1)^j [\det \mathbf{\Phi}_{(j)}(\mathbf{x})]. \tag{2.8}$$

Obviously, we have

$$d(s^n) = \sum_{j=0}^{n} d_{(j)}(s^n). \tag{2.9}$$

The value $d(s^n)$ may be positive or negative. The simplex $s^n$ is said to have a positive or negative orientation relative to the $x_i$ coordinate system according to whether $d(s^n) > 0$ or $< 0$. Of course, we also have

$$|d(s^n)| = n! \,(\text{the volume of } |s^n|) \tag{2.10}$$

where, as usual, $|s^n|$ stands for the point set of $s^n$ in $R^n$. Here we alert the reader that the use of vertical bars to indicate a simplex as a point set will be sometimes dispensed with when there is no possibility of ambiguity.

Next, consider a vertex $\mathbf{x}_j$ from the vertex set $V$. Its complement set $\{\mathbf{x}_j\}^c$ defines an $(n-1)$-simples which will be denoted as $s_{(j)}^{n-1}$:

$$s_{(j)}^{n-1} = \Delta(\mathbf{x}_0, \mathbf{x}_1, \ldots, \hat{\mathbf{x}}_j, \ldots, \mathbf{x}_n). \tag{2.11}$$

Often, it is convenient to say that the vertex $\mathbf{x}_j$ and the simplex $s_{(j)}^{n-1}$ are opposite to each other. The boundary of $s^n$ can be expressed in terms of $s_{(j)}^{n-1}$ as follows:

$$\partial s^n = \sum_{j=0}^{n} (-1)^j s_{(j)}^{n-1}. \tag{2.12}$$

In the subsequent section we shall also use the following notation. Given a set of points $S$ in $R^n$, the hyperplane of the smallest dimension $r$ which contains $S$ will be denoted by $\Lambda(S)$. It will be written as $\Lambda^r(S)$ if it is desirable to indicate the dimension. The convex hull of $S$ will be denoted by $\Gamma(S)$ or $\Gamma^r(S)$ where $r$ is the dimension of the hull.

**3. Barycentric coordinates.** Since the barycentric coordinates will play a key role in the following development, we single out this topic for a discussion in this section. Let $V$ be a set of $(n+1)$ independent points $\mathbf{x}_i$, $i = 0, 1, \ldots, n$, in $R^n$. Then any point $\mathbf{x}$ in $R^n$ may be expressed as

$$\mathbf{x} = \sum_{i=0}^{n} t_i \mathbf{x}_i, \qquad \sum_{i=0}^{n} t_i = 1. \tag{3.1}$$

The coefficients $t_i$ are the barycentric coordinates of $\mathbf{x}$ relative to $V$ [8].

The set $V$ defines a simples $s^n$. The barycentric coordinates of a point characterize the location of the point relative to $s^n$ in a very attractive way. We enumerate below various results. For the ease of future reference the listing is perhaps more detailed than absolutely necessary.

(I) If $t_j > 0$, then $\mathbf{x}_j$ and $\mathbf{x}$ are on the same side of $\Lambda^{n-1}(s_{(j)}^{n-1})$. Here, since the meaning is clear, we use $s_{(j)}^{n-1}$ instead of $|s_{(j)}^{n-1}|$ to denote the point set.

(II) If $t_j < 0$, then $\mathbf{x}_j$ and $\mathbf{x}$ are on the opposite sides of $\Lambda^{n-1}(s_{(j)}^{n-1})$.

(III) If $t_j = 0$, then $\mathbf{x}$ is in $\Lambda^{n-1}(s_{(j)}^{n-1})$.

(IV) If $t_i > 0$ for all $i$, then $\mathbf{x}$ is in $s^n$. the converse is also true.

(V) If one of the $t_i$'s is negative, then $\mathbf{x}$ is outside $\bar{s}^n$, the closure of $s^n$.

(VI) If $t_j = 0$ and all other $t_i$'s are positive, then $\mathbf{x}$ is in $s_{(j)}^{n-1}$.

(VII) If $t_j = 0$ and $t_k = 0, j \neq k$, then $\mathbf{x}$ is in $\Lambda^{n-2}(\{\mathbf{x}_j, \mathbf{x}_k\}^c)$.

(VIII) If $t_j = 0$ and $t_k = 0$, $j < k$, and all other $t_i$'s are positive, then $\mathbf{x}$ is in $s_{(j,k)}^{n-2} = \Delta(\mathbf{x}_0, \mathbf{x}_1, \ldots, \hat{\mathbf{x}}_j, \ldots, \hat{\mathbf{x}}_k, \ldots, \mathbf{x}_n)$.

(IX) If $t_{j_1} = t_{j_2} = \cdots = t_{j_r} = 0$, then $\mathbf{x}$ is in $\Lambda^{n-r}(\{\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_r}\}^c)$.

(X) If $t_{j_1} = t_{j_2} = \cdots = t_{j_r} = 0$ and all other $t_i$'s are positive, then $\mathbf{x}$ is in

$$s_{(j_1, j_2, \ldots, j_r)}^{n-r} = \Delta\left(\{\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_r}\}^c\right).$$

(XI) If $t_{j_1} = t_{j_2} = \cdots = t_{j_n} = 0$, then the remaining barycentric coordinate, say $t_J$, is necessarily one and $\mathbf{x}$ is at $\mathbf{x}_J$.

**4. Triangulation of an $n$-cube into $n$-simplexes.** In a euclidean $n$-space the simplest and the most easily defined building blocks are undoubtedly the rectangular parallelepipeds. On the other hand, as well be seen in Secs. 5 and 6, for an $n$-dimensional problem the $n$-simplexes are the basic geometric units over which we can carry out a rigorous search procedure for the presence of a zero of a vector function. For this reason it is necessary to examine how an $n$-dimesional rectangular parallelepiped can be partitioned into a number of $n$-simplexes.

Since we are only interested in the manner by which such a partitioning can be accomplished, it is obvious that, instead of an $n$-dimensional rectangular parallelepiped, we can consider a unit $n$-cube in $R^n$. Such an $n$-cube has $2^n$ vertices. Without a loss of generality we take them to be a

$$x_i = 0 \quad \text{or} \quad 1, \qquad i = 1, 2, \ldots, n. \tag{4.1}$$

First, we shall describe three different ways by which the $2^n$ vertices will be designated.

(i) The vertices are designated as $p_{\delta_1,\delta_2\cdots\delta_n}$ where

$$\delta_i = \begin{cases} 0 & \text{if } x_i = 0, \\ 1 & \text{if } x_i = 1, \end{cases} \qquad i = 1,2,\ldots,n. \tag{4.2}$$

(ii) Sometimes $p_{\delta_1,\delta_2\cdots\delta_j}$, $1 \leqslant j < n$, will be used. Such a notation always implies that $\delta_i = 0$ for $j < i \leqslant n$. Therefore, it actually stands for $p_{\delta_1\delta_2\cdots\delta_j 00\cdots 0}$. With this notation we can write $p_{00\cdots 0}$ as $p_0$, $p_{100\cdots 0}$ as $p_1$, $p_{010\cdots 0}$ as $p_{01}$, etc..

(iii) The quantity $\delta_1\delta_2\cdots\delta_n$ can also be taken as a binary number with $\delta_1$ being the first binary digit, $\delta_2$ the second, and so forth. Thus the totality of $\delta_1\delta_2\cdots\delta_n$ can be made to correspond, in a one-to-one fashion, to the numbers from 0 to $2^n - 1$. In other words, we can denote

$$\begin{array}{llll} p_{000\cdots 0} & \text{as} & p_0, & p_{100\cdots 0} \quad \text{as} \quad p_1, \\ p_{010\cdots 0} & \text{as} & p_2, & \vdots \qquad\qquad \vdots \\ \vdots & & \vdots & \\ & & & p_{111\cdots 1} \quad \text{as} \quad p_{2^n-1}. \end{array} \tag{4.3}$$

In Fig. 1 an example of this vertex labelling is shown for a 3-cube.

Each notation has its advantages and disadvantages. (i) and (ii) are more informative but (iii) is more convenient. In the following discussion we shall use the one which is the most appropriate in a given circumstance. The symbols $p_0$ and $p_1$ appear in both (ii) and (iii). However, since they represent the same vertices in either schemes, there could be no possibility of confusion.

*Vertex set of the cube.* The complete set of $2^n$ vertices will be called the vertex set of the $n$-cube and denoted by $P$:

$$P = \{p_0, p_1, \ldots, p_{2^n-1}\}. \tag{4.4}$$

*A simplex and its ordered vertices.* Let $P_1$ be a subset of $P$ with members $p_{j_0}, p_{j_1}, \ldots, p_{j_r}$. If these vertices are independent, they define a $r$-simplex in $R^n$. Again as in Sec. 2, we take this simplex to be so defined that the vertices are ordered in accordance with the magnitudes of their binary designation numbers of (4.3).

In the following discussion the notation $R^j(x_1x_2\cdots x_j)$ will denote the $j$-dimensional hyperplane determined by

$$x_i = 0, \qquad i = j+1,\ j+2,\ldots,n. \tag{4.5}$$

We are now ready to describe the partitioning of an $n$-cube [8]. We proceed in an inductive manner starting with 1-cubes. The constructive analysis leads to a very simple and easily implementable scheme of partitioning for a cube of any dimension.

*1-cubes.* Consider first the hyperplane $R^1(x_1)$. This hyperplane of dimension 1 is, in fact, the straight line of the $x_1$-axis. Two members $p_0$ and $p_1$ of $P$ lie in this hyperplane. The line segment $p_0p_1$ is a 1-cube in $R^1(x_1)$. On the other hand, it is also a 1-simplex. Hence, in the one-dimensional case a 1-cube is a 1-simplex and there is no need of partitioning. This particular 1-simplex will be designated as $s_1^1(x_1)$ where $(x_1)$ is used to emphasize the fact that this simplex is in $R^1(x_1)$, and the subscript 1 is merely an identification number.

2-*cubes*. Next, we consider a 2-cube which is constructed by sweeping $s_1^1(x_1)$ in the $x_2$-direction by a distance of one unit. This creates a 2-cube by bringing two additional vertices $p_{01}$ and $p_{11}$ (or $p_2$ and $p_3$) into the picture. We note that $p_{01}$ and $p_{11}$ are, respectively, the end points of the sweep for $p_0$ and $p_1$. The construction also creates three 1-simplexes $\Delta(p_0, p_{01})$, $\Delta(p_1, p_{11})$, and $\Delta(p_{01}, p_{11})$. Obviously, this 2-cube in $R^2(x_1x_2)$ can be partitioned into two 2-simplexes. In fact, there are two ways of doing this. We shall adopt one of them. Among the new vertices we take the first one, in this case $p_{01}$. We post-adjoin $p_{01}$ with $s_1^1(x_1)$ to create the first 2-simplex $s_1^2(x_1x_2)$,

$$s_1^2(x_1x_2) = \Delta(p_0, p_1, p_{01}), \tag{4.6}$$

where the $(x_1x_2)$ designation is used to emphasize that this is a 2-simplex in $R^2(x_1x_2)$. This 2-simplex creates a new boundary 1-simplex $\Delta(p_1, p_{01})$. Next, we take the second new vertex $p_{11}$ and post-adjoin it with $\Delta(p_1, p_{01})$ to create the second 2-simplex $s_2^2(x_1x_2)$,

$$s_2^2(x_1x_2) = \Delta(p_1, p_{01}, p_{11}). \tag{4.7}$$

Since these two 2-simplexes exhaust the 2-cube in $R^2(x_1x_2)$, the partitioning process is completed. The 2-cube may also be regarded as an integral 2-chain with $s_1^2(x_1x_2)$ and $s_2^2(x_1x_2)$ as the bases [3]. We have

$$C^2 = s_1^2(x_1x_2) - s_2^2(x_1x_2). \tag{4.8}$$

The boundary of this 2-chain is given by

$$\partial C^2 = \Delta(p_0, p_1) - \Delta(p_0, p_{01}) + \Delta(p_1, p_{11}) - \Delta(p_{01}, p_{11}). \tag{4.9}$$

In (4.9) the 1-simplex $\Delta(p_1, p_{01})$ is absent because it appears as a boundary simplex twice in opposite senses. In fact, the integer coefficients $+1$ and $-1$ in (4.8) are determined on the condition that this interior 1-simplex should have opposite orientations when it appears as a boundary simplex to the two neighboring 2-simplexes.

3-*cubes*. To create a 3-cube we sweep the 2-cube in the $x_3$-direction by a distance of one unit. This brings into the picture four new vertices $p_{001}$, $p_{101}$, $p_{011}$, and $p_{111}$ which are, respectively, the end points of the sweep for $p_{00}$, $p_{10}$, $p_{01}$, and $p_{11}$. The partition of the 2-cube in the $R^2(x_1x_2)$ plane into two 2-simplexes $s_1^2(x_1x_2)$ and $s_2^2(x_1x_2)$ induces naturally a preliminary partition of the 3-cube into two prisms which have $s_1^2(x_1x_2)$ and $s_2^2(x_1x_2)$ as their bases and a height of a unit length in the $x_3$-drection, as seen in Fig. 1. Therefore, we need only to consider partitioning each prism into appropriate 3-simplexes. To create these 3-simplexes we describe below a procedure which can readily be extended to cubes of higher dimensions. Take the first prism with the base $s_1^2(x_1x_2)$ and its three new vertices $p_{001}$, $p_{101}$, and $p_{011}$.

(i) We post-adjoin the first new vertex $p_{001}$ (or $p_4$) to $s_1^2(x_1x_2)$ to form a 3-simplex which is $\Delta(p_0, p_1, p_2, p_4)$ and will be denoted by $s_1^3(x_1x_2x_3)$. In creating this 3-simplex, a new interior 2-simplex $\Delta(p_1, p_2, p_4)$ is introduced.

(ii) We then take the next new vertex $p_{101}$ (or $p_5$) and post-adjoin it to that new interior 2-simplex $\Delta(p_1, p_2, p_4)$. This creates a new 3-simplex $\Delta(p_1, p_2, p_4, p_5)$ which will be denoted as $s_2^3(x_1x_2x_3)$. In creating this 3-simplex, a new interior 2-simplex $\Delta(p_2, p_4, p_5)$ is introduced.

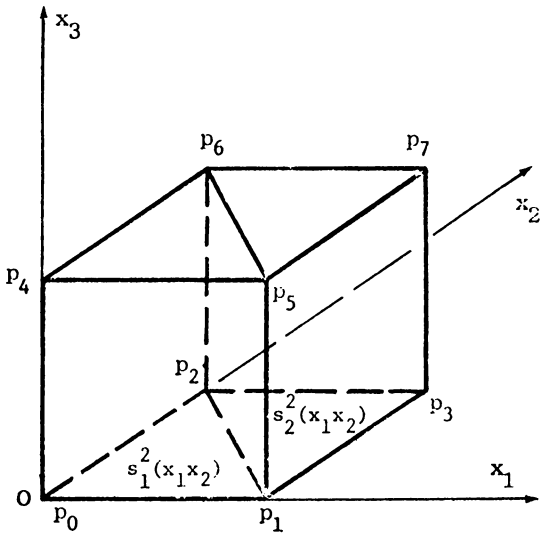


FIG. 1. Designation of the vertices of a 3-cube.

(iii) Next, we take the third new vertex $p_{011}$ (or $p_6$) and post-adjoin it to the new interior 2-simplex $\Delta(p_2, p_4, p_5)$ introduced in the preceding step. This creates a third 3-simplex $\Delta(p_2, p_4, p_5, p_6)$ which will be denoted as $s_3^3(x_1x_2x_3)$. One can then show that these three 3-simplexes exhaust the first prism.

The second prism having $s_2^2(x_1x_2)$ as its base and three new vertices $p_{101}$, $p_{011}$, and $p_{111}$ can be partitioned in exactly the same way, leading to three additional 3-simplexes which will be denoted as $s_4^3(x_1x_2x_3)$, $s_5^3(x_1x_2x_3)$, and $s_6^3(x_1x_2x_3)$. Thus, this mode of triangulation of a 3-cube yields the following six 3-simplexes:

$$\begin{aligned} s_1^3(x_1x_2x_3) &= \Delta(p_0, p_1, p_2, p_4), & s_2^3(x_1x_2x_3) &= \Delta(p_1, p_2, p_4, p_5), \\ s_3^3(x_1x_2x_3) &= \Delta(p_2, p_4, p_5, p_6), & s_4^3(x_1x_2x_3) &= \Delta(p_1, p_2, p_3, p_5), \\ s_5^3(x_1x_2x_3) &= \Delta(p_2, p_3, p_5, p_6), & s_6^3(x_1x_2x_3) &= \Delta(p_3, p_5, p_6, p_7). \end{aligned} \tag{4.10}$$

As an integral 3-chain with bases $s_i^3(x_1x_2x_3)$, $i = 1, 2, \ldots, 6$, the 3-cube is given by

$$C^3 = s_1^3 + s_2^3 + s_3^3 - s_4^3 - s_5^3 - s_6^3 \tag{4.11}$$

where the supplementary designation $(x_1x_2x_3)$ has been dropped. The boundary of this 3-chain is given by

$$\begin{aligned} \partial C^3 = &-\Delta(p_0, p_1, p_2) + \Delta(p_0, p_1, p_4) - \Delta(p_0, p_2, p_4) + \Delta(p_1, p_2, p_3) \\ &+\Delta(p_1, p_3, p_5) - \Delta(p_1, p_4, p_5) - \Delta(p_2, p_3, p_6) + \Delta(p_2, p_4, p_6) \\ &-\Delta(p_3, p_5, p_7) + \Delta(p_3, p_6, p_7) + \Delta(p_4, p_5, p_6) - \Delta(p_5, p_6, p_7). \end{aligned} \tag{4.12}$$

*4-cubes.* To create a 4-cube we sweep the 3-cube of $R^3(x_1x_2x_3)$ in the $x_4$-direction by a unit distance. This brings in eight new vertices. The six 3-simplexes of the 3-cube generate six "prisms". Each prism has four new vertices and it can be partitioned into four 4-simplexes. Thus, the 4-cube is partitioned into altogether twenty-four 4-simplexes $s_i^4(x_1x_2x_3x_4)$, $i = 1, 2, \ldots, 24$. Without the $(x_1x_2x_3x_4)$ designation and the separating

commas inside $\Delta(\cdots)$, they are

$$
\begin{aligned}
s_1^4 &= \Delta(p_0p_1p_2p_4p_8), & s_2^4 &= \Delta(p_1p_2p_4p_8p_9),\\
s_3^4 &= \Delta(p_2p_4p_8p_9p_{10}), & s_4^4 &= \Delta(p_4p_8p_9p_{10}p_{12}),\\
s_5^4 &= \Delta(p_1p_2p_4p_5p_9), & s_6^4 &= \Delta(p_2p_4p_5p_9p_{10}),\\
s_7^4 &= \Delta(p_4p_5p_9p_{10}p_{12}), & s_8^4 &= \Delta(p_5p_9p_{10}p_{12}p_{13}),\\
s_9^4 &= \Delta(p_2p_4p_5p_6p_{10}), & s_{10}^4 &= \Delta(p_4p_5p_6p_{10}p_{12}),\\
s_{11}^4 &= \Delta(p_5p_6p_{10}p_{12}p_{13}), & s_{12}^4 &= \Delta(p_6p_{10}p_{12}p_{13}p_{14}),\\
s_{13}^4 &= \Delta(p_1p_2p_3p_5p_9), & s_{14}^4 &= \Delta(p_2p_3p_5p_9p_{10}),\\
s_{15}^4 &= \Delta(p_3p_5p_9p_{10}p_{11}), & s_{16}^4 &= \Delta(p_5p_9p_{10}p_{11}p_{13}),\\
s_{17}^4 &= \Delta(p_2p_3p_5p_6p_{10}), & s_{18}^4 &= \Delta(p_3p_5p_6p_{10}p_{11}),\\
s_{19}^4 &= \Delta(p_5p_6p_{10}p_{11}p_{13}), & s_{20}^4 &= \Delta(p_6p_{10}p_{11}p_{13}p_{14}),\\
s_{21}^4 &= \Delta(p_3p_5p_6p_7p_{11}), & s_{22}^4 &= \Delta(p_5p_6p_7p_{11}p_{13}),\\
s_{23}^4 &= \Delta(p_6p_7p_{11}p_{13}p_{14}), & s_{24}^4 &= \Delta(p_7p_{11}p_{13}p_{14}p_{15}).
\end{aligned}
\tag{4.13}
$$

As an integral 4-chain the 4-cube is given by

$$
\begin{aligned}
C^4 = {} & s_1^4 - s_2^4 + s_3^4 - s_4^4 + s_5^4 - s_6^4 + s_7^4 - s_8^4 + s_9^4 - s_{10}^4 + s_{11}^4 - s_{12}^4\\
& - s_{13}^4 + s_{14}^4 - s_{15}^4 + s_{16}^4 - s_{17}^4 + s_{18}^4 - s_{19}^4 + s_{20}^4 - s_{21}^4 + s_{22}^4 - s_{23}^4 + s_{24}^4.
\end{aligned}
\tag{4.14}
$$

The boundary $\partial C^4$ can be expressed in terms of forty eight 3-simplexes.

*n-cubes.* The proposed method of partitioning an $n$-cube into $n$-simplexes is now clear. The total number of the partitioned $n$-simplexes is $n!$. The complete list of the simplexes can easily be written down when needed. The number of $(n-1)$-simplexes making up the boundary of an $n$-cube is $2(n!)$.

**5. Linear functions and simplical mappings.** In Secs. 2–4 we have presented the necessary geometrical framework for the proposed method of locating the zeros of a function. In this and the next sections we discuss how to use an approximating linear function over an $n$-simplex to test the possible presence of a zero of a general nonlinear function in the simplex. The development leads eventually to an operational description of the method in Secs. 7 and 8.

Consider a real valued linear function $f^L$: $X^n \to F^n$ given in the form

$$
\mathbf{f}^L(\mathbf{x}) = \mathbf{A}_0 + \mathbf{A}\mathbf{x}, \qquad \mathbf{x} \in X^n, \mathbf{f}^L \in F^n, \tag{5.1}
$$

where $\mathbf{A}_0$ is an $n$-vector, $\mathbf{A}$ is an $n \times n$ matrix, and $X^n$ and $F^n$ are two euclidean $n$-spaces. Here we use the superscript $L$ to indicate the linearity of the function. This notation is useful because it allows us later to use $\mathbf{f}^L(\mathbf{x})$ to represent a local linear approximation to a given, and generally nonlinear, function $\mathbf{f}(\mathbf{x})$. However, since in this and the next sections our analysis will be almost entirely confined to linear functions like (5.1), we shall drop the superscript $L$ temporarily.

We first consider in this section the nondegenerate case where $(\det \mathbf{A}) \neq 0$. For this case we can write

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}(\mathbf{x} - \mathbf{x}^*) \quad \text{with } \mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{A}_0. \tag{5.2}$$

Evidently, $\mathbf{x}^*$ is the zero of $\mathbf{f}(\mathbf{x})$. Consider now an $n$-simplex in $X^n$ with a vertex set $V_X$ consisting of vertices $\mathbf{x}_j, j = 0, 1, \ldots, n$. The function $\mathbf{f}(\mathbf{x})$ maps these vertices into $(n + 1)$ points in $F^n$. These points will be simply denoted as $\mathbf{f}_j, j = 0, 1, \ldots, n$, with $\mathbf{f}_j = \mathbf{f}(\mathbf{x}_j)$. They form a vertex set $V_F$ of an $n$-simplex $\sigma^n$ in $F^n$. Thus, the function $\mathbf{f}(\mathbf{x})$ defines a simplicial mapping; in fact, a nondegenerate one when $\mathbf{f}(\mathbf{x})$ is nondegenerate [3]. Here and in the following sections we shall always use $\sigma^r$ to denote a $r$-simplex in $F^n$ while keeping $s^r$ to denote a $r$-simplex in $X^n$. In this context of a simplicial mapping our goal may be restated as one to find a simple characterization of the location of the zero $x^*$ relative to $s^n$, based upon the information on $\sigma^n$. We shall discuss the matter from four different points of view: geometrical, algebraic, search testing, and programmatic.

(I) *Geometrical point of view.* By definition $\mathbf{x}^*$ is mapped into $O_F$, the origin of $F^n$. Since a nondegenerate linear mapping maps the interior of a simplex into the interior of its image simplex and the exterior to the exterior, one has the following results.

THEOREM 5.1. Let $\mathbf{f}(\mathbf{x})$ be given by (5.1) and $(\det \mathbf{A}) \neq 0$. Let $s^k$ be a $k$-simplex in $X^n$ and $\sigma^k$ be the corresponding image simplex of $s^k$ in $F^n$, $0 \leqslant k \leqslant n$. The necessary and sufficient condition for $\mathbf{x}^*$ of $\mathbf{f}(\mathbf{x})$ to be inside $s^k$ is that $\sigma^k$ contains $O_F$ in its interior.

The results of Theorem 5.1 are almost self-evident and trivial and yet they lead us to some remarkable consequences. For our purpose the most important case is when $k = n$. The zero $x^*$ is inside $s^n$ if and only if $\sigma^n$ contains $O_F$ in its interior.

(II) *Algebraic point of view.* When $\mathbf{f}(\mathbf{x})$ of (5.1) is given, we can easily compute the values of $\mathbf{f}(\mathbf{x})$ at the vertices of $s^n$ in $X^n$. Conversely, if $(n + 1)$ function values at the vertices of $s^n$ are known, they uniquely determine a linear function $\mathbf{f}(\mathbf{x})$, degenerate or nondegenerate. This means $\mathbf{A}_0$ and $\mathbf{A}$ can be determined in terms of $\mathbf{f}_j, j = 0, 1, \ldots, n$. When $\mathbf{f}(\mathbf{x})$ thus determined is nondegenerate, its unique zero can then be evaluated by using the second equation of (5.2) and its location relative to the simplex $s^n$ can be examined accordingly. In the following we shall, however, take an entirely different approach leading to certain analytic results which can serve as the basis of an effective searching algorithm for zeros.

First let us define $\mathbf{x}^+$ and $\mathbf{f}^+$ as an augmented $\mathbf{x}$ vector and an augmented $\mathbf{f}$ vector according to

$$\mathbf{x}^+ = \begin{bmatrix} 1 \\ \mathbf{x} \end{bmatrix}, \qquad \mathbf{f}^+ = \begin{bmatrix} 1 \\ \mathbf{f} \end{bmatrix}. \tag{5.3}$$

Next, we define a new $(n + 1) \times (n + 1)$ matric $\mathbf{A}^+$ by

$$\mathbf{A}^+ = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{A}_0 & \mathbf{A} \end{bmatrix}. \tag{5.4}$$

Evidently $\mathbf{A}^+$ is nonsingular if and only if $\mathbf{A}$ is nonsingular. In terms of $\mathbf{x}^+$ and $\mathbf{f}^+$, (5.1) can be written as

$$\mathbf{f}^+(\mathbf{x}) = \mathbf{A}^+ \mathbf{x}^+ . \tag{5.5}$$

It is seen here that $\mathbf{A}^+$ is the matrix representation of an affine mapping.

Applying (5.5) to the $(n+1)$ vertices of $V_X$ and arranging the two sets augmented vectors $\mathbf{x}_i^+$ and $\mathbf{f}_i^+$, $i = 0, 1, \ldots, n$, in the form of (2.6), we get

$$\mathbf{\Phi}^+(\mathbf{f}) = \mathbf{A}^+ \mathbf{\Phi}^+(\mathbf{x}). \tag{5.6}$$

We may interpret $\mathbf{\Phi}^+(\mathbf{x})$ and $\mathbf{\Phi}^+(\mathbf{f})$ as the matrix representations of the vertex sets $V_X$ and $V_F$. (5.6) can then be regarded as the mapping from $V_X$ to $V_F$ or from $s^n$ to $\sigma^n$. From (5.6) we also have

$$d(\sigma^n) \equiv \det \mathbf{\Phi}^+(\mathbf{f}) = (\det \mathbf{A})(\det \mathbf{\Phi}^+(\mathbf{x})) \equiv (\det \mathbf{A}) d(s^n) \tag{5.7}$$

$$(\text{volume of } |\sigma^n|) = |(\det A)| (\text{volume of } |s^n|). \tag{5.8}$$

Let $\mathbf{t}$ be a column vector with the barycentric coordinates as its components,

$$\mathbf{t} = [t_0, t_1, \ldots, t_n]^T. \tag{5.9}$$

Then by (3.1) and (5.3) we have

$$\mathbf{x}^+ = \mathbf{\Phi}^+(\mathbf{x})\mathbf{t}. \tag{5.10}$$

Using (5.10) and (5.6), we can express $\mathbf{f}^+(\mathbf{x})$ of (5.5) in the form

$$\mathbf{f}^+(\mathbf{x}) = \mathbf{\Phi}^+(\mathbf{f})\mathbf{t}. \tag{5.11}$$

This will be seen to be the key equation of our development. It maps an arbitrary point in $X^n$ with barycentric coordinates $t_i$, $i = 0, 1, \ldots, n$, to a point in $F^n$. Taking the inverse of (5.11), we obtain

$$\mathbf{t} = [\mathbf{\Phi}^+(\mathbf{f})]^{-1} \mathbf{f}^+(\mathbf{x}). \tag{5.12}$$

The origin of $F^n$ is given by $\mathbf{f}^+ = [1, 0, 0, \ldots, 0]^T$. Using this in (5.12), we readily find the barycentric coordinates of the zero $\mathbf{x}^*$ of $\mathbf{f}(\mathbf{x})$ given by

$$t_j = d_{(j)}(\sigma^n)/d(\sigma^n), \qquad j = 0, 1, \ldots, n, \tag{5.13}$$

where

$$d_{(j)}(\sigma^n) = (-1)^j \det \mathbf{\Phi}_{(j)}(\mathbf{f}) \tag{5.14}$$

$$d(\sigma^n) = \det \mathbf{\Phi}^+(\mathbf{f}) = \sum_{j=0}^{n} d_{(j)}(\sigma^n). \tag{5.15}$$

Thus, given a set of $(n+1)$ function values at the vertex set $V_X$, $(n+1)$ determinants $d_{(j)}(\sigma^n)$ of (5.14) can be evaluated. They determine the barycentric coordinates of the zero $\mathbf{x}^*$ of $\mathbf{f}(\mathbf{x})$ in $X^n$. Once the barycentric coordinates are known, the full force of results (I)–(XI) discussed in Sec. 3 can be brought to bear to examine the location of $\mathbf{x}^*$ relative to the simplex $s^n$.

(III) *Search testing point of view.* In principle, (5.13) is the only equation we need to determine the location of $\mathbf{x}^*$. However, if our aim is merely to ascertain whether the zero lies inside $s^n$ or on one of its proper faces, then actually computing $\mathbf{x}^*$ is not a desirable way to proceed. Other kinds of testing procedures are much more preferable. We describe here one such procedure which is based upon comparing the signs of the $(n+1)$ determinats $d_{(i)}(\sigma^n)$. Since the degenerate case will be discussed separately in the next section, we assume $d(\sigma^n) \neq 0$. For convenience we refer to $d_{(i)}(\sigma^n)$ as the *ith basic F-determinant* and denote it by an abbreviated notation $d_{(i)}^F$.

(A) None of the $(n + 1)$ basic $F$-determinants vanish:

(Ai) If there are two $d^F_{(j)}$ and $d^F_{(k)}$ which are opposite in sign, then there is at least one negative barycentric coordinate. By result (V) of Sec. 3 the zero is outside $\bar{s}^n$.

(Aii) If the $(n + 1)$ basic $F$-determinants are all of the same sign, then by result (IV) the zero is inside $s^n$.

(B) One and only one basic $F$-determinant vansihes. Let $d^F_{(j)} = 0$. This implies $t_j = 0$ and, therefore, by result (III) the zero lies in $\Lambda^{n-1}(s^{n-1}_{(j)})$.

(Bi) If, among the other $n$ basic $F$-determinants, there are two which are opposite in sign, then there is again a negative barycentric coordinate and, by result (V), the zero is outside $\bar{s}^n$.

(Bii) If all the other $n$ basic $F$-determinants are all of the same sign, then by result (VI) the zero is in $s^{n-1}_{(j)}$.

(C) There are $r$, and only $r$, basic $F$-determinants equal to zero. Let them be $d^F_{(j_1)}$, $d^F_{(j_2)}, \ldots, d^F_{(j_r)}$. In this case $t_{j_1} = t_{j_2} = \cdots = t_{j_r} = 0$ and, therefore, by result (IX) the zero lies in $\Lambda^{n-r}(\{\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_r}\}^c)$.

(Ci) If, among the other $n - r + 1$ nonvanishing $F$-determinants, there are two which are opposite in sign, then there is again a negative barycentric coordinate and the zero lies outside $\bar{s}^n$.

(Cii) If all the $n - r + 1$ remaining nonvanishing $F$-determinants are all of the same sign, then by result $(X)$ the zero lies inside the proper face

$$s^{n-r}_{(j_1, j_2, \ldots, j_r)} = \Delta\left(\{\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_r}\}^c\right).$$

(D) All the basic $F$-determinants except two vanish. Let the nonvanishing ones be $d^F_{(j)}$ and $d^F_{(k)}$, $j < k$. Then $t_j$ and $t_k$ are different from zero and the zero $\mathbf{x}^*$ lies in the line $\Lambda^1(\mathbf{x}_j, \mathbf{x}_k)$. In this case it can be shown that $\mathbf{f}_j$ and $\mathbf{f}_k$ are necessarily parallel to each other.

(Di) If $d^F_{(j)}$ and $d^F_{(k)}$ are opposite in sign, then one of $t_j$ and $t_k$ is negative and the zero is outside $\bar{s}^n$ by result (V). In this case one can show that $\mathbf{f}_j$ and $\mathbf{f}_k$ are in the same direction.

(Dii) If $d_{(j)}$ and $d_{(k)}$ are equal in sign, then by result (X) the zero is in the 1-simplex $\Delta(\mathbf{x}_j, \mathbf{x}_k)$. In this case $\mathbf{f}_j$ and $\mathbf{f}_k$ are opposite in direction.

(E) All the basic $F$-determinants except $d^F_{(j)}$ vanish. In this case only $t_j$ is nonzero and equal to 1. The zero is at the vertex $\mathbf{x}_j$.

The above discussion provides us with a rigorous basis for a testing procedure to determine whether the zero of a nondegenerate linear function is inside or on the boundary of an $n$-simplex. The testing is entirely done by examining the signs and the vanishing of the basic $F$-determinants. Once it has been ascertained that the zero does lie in the closure of the simplex we can compute $\mathbf{x}^*$ by (5.13) and (3.1), if so desired.

(IV) *Programmatic point of view.* The above discussion gives us a very clear picture about how the location of $\mathbf{x}^*$ relative to $s^n$ depends upon the $(n + 1)$ basic $F$-determinants. When we search test a large number of $n$-simplexes for zeros, most of the simplexes will contain no zeros. It is quite likely that the comparing of the first two or three determinants will already show that there is no zero inside the simplex. For these cases, to compute all the basic $F$-determinants first and then examine their signs will be a very wasteful procedure. In devising an implementing algorithm, we should, therefore, alternate

the determinant computation and the sign comparison so that once two basic $F$-determinants of opposite signs have been discovered, the search testing can be terminated immediately for that simplex.

**6. Degenerate linear functions.** In this section we study the degenerate case where **A** of (5.1) is singular. In that case $d(\sigma^n) = 0$ by (5.7) and [Rank $\mathbf{\Phi}^+(\mathbf{f})$] $< n + 1$. Therefore, the image simplex of $s^n$ collapses into a complex of dimension $r$ less than $n$. Let $V_F$ again denote the image set of the vertex set $V_X$ of $s^n$. Let $\Gamma^r(V_F)$ be the convex hull of $V_F$ and let $\Lambda^r(V_F)$ be the smallest hyperplane containing $V_F$.

First we shall make a general observation concerning the existence of zeros of $\mathbf{f}(\mathbf{x})$. Since the mapping is affine, the complete space $X^n$ is mapped into $\Lambda^r(V_F)$ of $F^n$. It then follows that if $\Lambda^r(V_F)$ does not contain $O_F$, the origin of $F^n$, there exists no zero of $\mathbf{f}(\mathbf{x})$ in $X^n$. For convenience we refer to this condition of $O_F$ not in $\Lambda^r(V_F)$ as the *no zero condition*. Moreover, we have the following result.

THEOREM 6.1. Let $\mathbf{f}(\mathbf{x})$ be an affine mapping from $X^n$ to $F^n$ as given by (4.1). Let $s^n$ be an $n$-simplex in $X^n$. The equation $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ has a solution if and only if

$$\text{Rank } \mathbf{\Phi}^+(\mathbf{f}) = \text{Rank } \mathbf{\Phi}(\mathbf{f}) + 1. \tag{6.1}$$

*Proof.* By (5.11) the equation $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ is equivalent to

$$\mathbf{\Phi}^+(\mathbf{f})\mathbf{t} = \mathbf{E}_0 \quad \text{where } \mathbf{E}_0 = [1, 0, 0, \ldots, 0]^T. \tag{6.2}$$

Let

$$\mathbf{\Phi}^{+\,+}(\mathbf{f}) = [\mathbf{\Phi}^+(\mathbf{f}), \mathbf{E}_0] \tag{6.3}$$

be an augmented matrix of $\mathbf{\Phi}^+(\mathbf{f})$ by $\mathbf{E}_0$. Then by the well-known theorem on nonhomogeneous systems of linear equations, (6.2) has a solution if and only if $\mathbf{\Phi}^+(\mathbf{f})$ and $\mathbf{\Phi}^{+\,+}(\mathbf{f})$ have the same rank. Since $\mathbf{\Phi}^{+\,+}(\mathbf{f})$ is *equivalent* to

$$\begin{bmatrix} \mathbf{0} & 1 \\ \mathbf{\Phi}(\mathbf{f}) & \mathbf{0} \end{bmatrix} \tag{6.4}$$

the conclusion follows immediagely.

Thus, the degenerate cases are separated by Theorem 6.1 into two groups: (i) (6.1) is not met. $\mathbf{f}(\mathbf{x})$ has no zeros. (ii) (6.1) is satisfied. $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ will have infinite number of solutions. For the second group it is also implied that $\Lambda^r(V_F)$ contains $O_F$. In what follows we study various cases of the second group in considerable detail. We note that if $r$ is the dimension of the smallest hyperplane containing the vertex set $V_F$, then the affine kernel space of the mapping [9], denoted by $K^q$ will have a dimension $q = n - r$. Another important question about the kernel space is whether it intersects with the simplex $s^n$. Again, qualitatively it is easy to answer. If $\Gamma^r(V_F)$ contains $O_F$, then $K^q$ and $s^n$ intersect. Actual testing of this fact is, however, not so simple. One also sees that if $O_F$ is on the boundary of $\Gamma^r(V_F)$, then the affine kernel subspace $K^q$ touches $s^n$ at one of its faces but does not penetrate into $s^n$ itself.

(A) $r = 0$. This is a trivial case. $\Gamma^0(V_F)$, the image of $s^n$, is a single point and, because of (6.1) it is in fact $O_F$. the function $\mathbf{f}(\mathbf{x})$ is identically zero, and Rank $\mathbf{\Phi}^+(\mathbf{f}) = 1$ and Rank $\mathbf{\Phi}(\mathbf{f}) = 0$. The kernel of the mapping is the whole space of $X^n$.

(B) $r = 1$. For this case Rank $\mathbf{\Phi}^+(\mathbf{f}) = 2$ and Rank $\mathbf{\Phi}(f) = 1$. $\Lambda^1(V_F)$ is a line passing through $O_F$. Therefore, all $\mathbf{f}_i$, $i = 0, 1, \ldots, n$, are parallel and we may write

$$\mathbf{f}_i = \alpha_i \mathbf{c} \tag{6.5}$$

where $\mathbf{c}$ is an $n$-vector in $F^n$. By (5.11) we have then

$$\mathbf{f}(\mathbf{x}) = \left\{ \sum_{i=0}^{n} \alpha_i t_i \right\} \mathbf{c} \tag{6.6}$$

and the kernel space of the mapping is given by

$$\sum_{i=0}^{n} \alpha_i t_i = 0, \qquad \sum_{i=0}^{n} t_i = 1. \tag{6.7}$$

The kernel space is of dimension $n - 1$.

With regard to the question of whether the affine kernel space $K^{n-1}$ intersects with $s^n$. We have the following results:

(i) If $\alpha_i$, $i = 0, 1, \ldots, n$, are not all of the same sign, then $O_F$ is in $\Gamma^1(V_F)$ and $K^{n-1}$ intersects with $s^n$.

(ii) If $\alpha_i$ are all of the same sign, $K^{n-1}$ does not intersect with $s^n$.

(iii) If $\alpha_{i_1} = \alpha_{i_2} = \cdots = \alpha_{i_k} = 0$, $k \leqslant n$, and the remaining $\alpha_i$ are all of the same sign, then the affine kernel space intersects with $s^n$ at the $(k-1)$-face $\Delta(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, h_{i_k})$.

(C) *The general case*. Rank $\mathbf{\Phi}^+(\mathbf{f}) = r + 1$ and Rank $\mathbf{\Phi}(\mathbf{f}) = r$. The subspace $\Lambda^r(V_F)$ spanned by $V_F$ is of dimension $r$. Take $\mathbf{c}_i$, $i = 1, 2, \ldots, r$, as a set of base vectors in this subspace and express $\mathbf{f}_j$ in terms of them

$$\mathbf{f}_j = \sum_{i=1}^{r} \mathbf{c}_i \alpha_{ij}, \qquad j = 0, 1, \ldots, n. \tag{6.8}$$

Using this in (5.11), we obtain

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{r} \left\{ \sum_{j=0}^{n} \alpha_{ij} t_j \right\} \mathbf{c}_i. \tag{6.9}$$

The affine kernel subspace $K^{n-r}$ is, therefore, determined by

$$\sum_{j=0}^{n} \alpha_{ij} t_j = 0, \qquad i = 1, 2, \ldots, r, \quad \sum_{j=0}^{n} t_j = 1. \tag{6.10}$$

To decide whether $K^{n-r}$ intersects with $s^n$ we can proceed as follows. $\Gamma^r(V_F)$, the image of $s^n$, is a convex polyhedron of dimension $r$. Let it be partitioned into a number of $r$-simplexes. We can then examine each $r$-simplex in turn. If a simplex is found to contain $O_f$, then it can be immediately concluded that the kernel space intersects with $s^n$. It may happen that $O_F$ is found to be on the boundary of a simplex which is also a part of the boundary of $\Gamma^r(V_F)$, then, as mentioned before, the kernel space touches $s^n$ at one of its faces but does not penetrate into $s^n$ itself.

Let $\sigma^r$ denote one of the partitioned $r$-simplexes. Let the vertices of $\sigma^r$ be at $\mathbf{f}_{k_0}$, $\mathbf{f}_{k_1}, \ldots, \mathbf{f}_{k_r}$, $k_0 < k_1 < \cdots < k_r$, to be denote by $V(\sigma^r)$ as a set. Any point $\mathbf{f}$ in $\Lambda^r(V_F)$ can

be expressed in terms of the barycentric coordinates $\tau_{k_\lambda}$ relative to this set of vertices

$$\mathbf{f} = \sum_{\lambda=0}^{r} \tau_{k_\lambda} \mathbf{f}_{k_\lambda}, \qquad \sum_{\lambda=0}^{r} \tau_{k_\lambda} = 1. \tag{6.11}$$

In terms of the base vectors $\mathbf{c}_i$, we have

$$\mathbf{f} = \sum_{i=1}^{r} \left( \sum_{\lambda=0}^{r} \alpha_{ik_\lambda} \tau_{k_\lambda} \right) \mathbf{c}_i. \tag{6.12}$$

We may also express $\mathbf{f}$ in $\Lambda^r(V_F)$ in terms of $\mathbf{c}_i$ directly

$$\mathbf{f} = \sum_{i=1}^{r} \alpha_i \mathbf{c}_i. \tag{6.13}$$

Equating (6.12) and (6.13), we obtain

$$\alpha_i = \sum_{\lambda=0}^{r} \alpha_{ik_\lambda} \tau_{k_\lambda}, \qquad \sum_{\lambda=0}^{r} \tau_{k_\lambda} = 1. \tag{6.14}$$

For further discussion it is convenient to introduce the following matrices:

$$\begin{aligned}
\boldsymbol{\alpha} &= [\alpha_1, \alpha_2, \ldots, \alpha_r]^T, \qquad \boldsymbol{\alpha}^+ = \begin{bmatrix} 1 \\ \boldsymbol{\alpha} \end{bmatrix}, \\
\boldsymbol{\alpha}_{k_\lambda} &= [\alpha_{1k_\lambda}, \alpha_{2k_\lambda}, \ldots, \alpha_{rk_\lambda}]^T, \qquad \boldsymbol{\alpha}^+_{k_\lambda} = \begin{bmatrix} 1 \\ \boldsymbol{\alpha}_{k_\lambda} \end{bmatrix}, \lambda = 0, 1, \ldots, r, \\
\boldsymbol{\tau} &= [\tau_{k_0}, \tau_{k_1}, \ldots, \tau_{k_r}]^T, \\
\boldsymbol{\Phi}(\boldsymbol{\alpha}) &= [\boldsymbol{\alpha}_{k_0}, \boldsymbol{\alpha}_{k_1}, \ldots, \boldsymbol{\alpha}_{k_r}], \qquad \boldsymbol{\Phi}^+(\boldsymbol{\alpha}) = \begin{bmatrix} 1,1,\ldots,1 \\ \boldsymbol{\alpha}_{k_0}, \boldsymbol{\alpha}_{k_1}, \ldots, \boldsymbol{\alpha}_{k_r} \end{bmatrix}.
\end{aligned} \tag{6.15}$$

The matrices $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}^+$, $\boldsymbol{\alpha}_{k_\lambda}$, $\boldsymbol{\alpha}^+_{k_\lambda}$, $\boldsymbol{\tau}$, $\boldsymbol{\Phi}(\boldsymbol{\alpha})$, and $\boldsymbol{\Phi}^+(\boldsymbol{\alpha})$ are, respectively, of the order $r \times 1$, $(r+1) \times 1$, $r \times 1$, $(r+1) \times 1$, $(r+1) \times 1$, $r \times (r+1)$, and $(r+1) \times (r+1)$. The equation (6.14) may now be written as

$$\boldsymbol{\alpha}^+ = \boldsymbol{\Phi}^+(\boldsymbol{\alpha}) \boldsymbol{\tau}. \tag{6.16}$$

Here, we note that $\boldsymbol{\Phi}^+(\boldsymbol{\alpha})$ is nonsingular.

The equation (6.16) has a very simple and useful interpretation. Let $A^r$ be a euclidean $r$-space with a coordinate system $\alpha_i$, $i = 1, 2, \ldots, r$. The set of $(r+1)$ $r$-vectors $\boldsymbol{\alpha}_{k_\lambda}$, $\lambda = 0, 1, \ldots, r$, of (6.15) may be taken as the vertex set $V_a$ of a $r$-simplex $\beta^r$ in $A^r$. If they are regarded as the image set of $V(\sigma^r)$, then (6.16) is the matrix representation of this mapping from a eulcidean $r$-space $\Lambda^r(V(\sigma^r))$ to a eulcidean $r$-space $A^r$. Moreover, this mapping is nondegenerate. As before, let us define

$$\begin{aligned}
\boldsymbol{\Phi}_{(\lambda)}(\boldsymbol{\alpha}) &= [\boldsymbol{\alpha}_{k_0}, \boldsymbol{\alpha}_{k_1}, \ldots, \hat{\boldsymbol{\alpha}}_{k_\lambda}, \ldots, \boldsymbol{\alpha}_{k_r}], \\
d^{\alpha}_{(\lambda)} &\equiv d_{(\lambda)}(\beta^r) = (-1)^\lambda [\det \boldsymbol{\Phi}_{(\lambda)}(\boldsymbol{\alpha})], \\
d(\beta^r) &\equiv \det \boldsymbol{\Phi}^+(\alpha) = \sum_{\lambda=0}^{r} d^{\alpha}_{(\lambda)}.
\end{aligned} \tag{6.17}$$

By (6.13) a zero $\mathbf{f}$ correspond to a zero $\boldsymbol{\alpha}$, or to an $\boldsymbol{\alpha}^+ = [1,0,0,\ldots,0]^T$. Using this in (6.16) and solving for $\tau$, we obtain the barycentric coordinates of $O_F$ relative to the simplex $\sigma^r$ in the hyperplane $\Lambda^r(V_F)$ as

$$\tau_{k_\lambda} = d^{\alpha}_{(\lambda)}/d(\beta^r), \qquad \lambda = 0,1,\ldots,r. \tag{6.18}$$

Once we have (6.18), the results (I)–(XI) of Sec. 3 may used to see whether $O_F$ is inside or on the boundary of $\sigma^r$. This treatment may be repeated for every partitioned $r$-simplex of $\Gamma^r(V_F)$, if necessary. The aggregated results then decide whether the affine kernel subspace $K^{n-r}$ intersects with $s^n$.

(D) $r = n - 1$. This case for which Rank $\boldsymbol{\Phi}^+(\mathbf{f}) = n$ and Rank $\boldsymbol{\Phi}(\mathbf{f}) = n - 1$ is, of course, covered by the general case. It is a degenerate case of the lowest degree in that the image of $s^n$ loses only one dimension under the mapping $\mathbf{f}(\mathbf{x})$. The affine kernel subspace is of dimension 1.

**7. First level search for zeros of a function.** We are now ready to discuss the proposed method of locating the zeros of a function. Consider an $n$-dimensional function $\mathbf{f}(\mathbf{x})$ of class $C^1$ over a euclidean $n$-space $X^n$. In general, the function is nonlinear. Let the task be to find *all* the zeros of this function in a domain of intersect defined by

$$x_i' \leqslant x_i \leqslant x_i'', \qquad i = 1,2,\ldots,n. \tag{7.1}$$

The proposed method is based upon the development given in Secs. 2–6. It consists of three main parts: setting up a structure of $n$-simplexes, first level search over individual simplexes, and refining the discovered zeros. We shall discuss the first two parts in this section and the third part in Sec. 8.

*Setting up a structure of n-simplexes.* First we choose a set of appropriate interval sizes $h_i$, $i = 1,2,\ldots,n$, in the $n$ coordinate directions

$$h_i = (x_i'' - x_i')/N_i, \qquad i = 1,2,\ldots,n, \tag{7.2}$$

where $N_i$ is the number of intervals in the $x_i$-direction. Using these interval sizes, we divide the domain of interest into $N$ $(= N_1 \times N_2 \times \cdots N_n)$ number of $n$-cubes, each having the size $h_1 \times h_2 \times \cdots \times h_n$. Each cube can, in turn, be partitioned into $n!$ $n$-simplexes according to the triangulation scheme presented in Sec. 4. In this manner a structure of $N(n!)$ $n$-simplexes is created. These simplexes are then arranged into a one-dimensional array with the members labelled from 1 to $N(n!)$. To each simplex in the array we apply the *First level search procedure* to see whether a potential zero of $\mathbf{f}(\mathbf{x})$ is present in this simplex. When the search through the array is completed, we expect to have located all the simple zeros of the function in the domain of interest, except perhaps those which are apart by a distance less than $h_i$ in the $x_i$-direction, $i = 1,2,\ldots,n$.

*First level search procedure over an n-simplex.* The first level search procedure for the presence of a zero is founded on the analysis given in Secs. 5 and 6. The basic idea is to approximate the given function $\mathbf{f}(\mathbf{x})$ at a particular $n$-simplex by a linear function $\mathbf{f}^L(\mathbf{x})$. If we *sample* the function $\mathbf{f}(\mathbf{x})$ at the vertices $\mathbf{x}_i$, $i = 0,1,\ldots,n$, of the $n$-simplex under search, we obtain $(n + 1)$ function values $\mathbf{f}_i$, $i = 0,1,\ldots,n$. These values determine a unique approximating function $\mathbf{f}^L(\mathbf{x})$. We can then use the development given in Sec. 5 to test

whether $\mathbf{f}^L(\mathbf{x})$ has its zero in the simplex. If $\mathbf{f}^L(\mathbf{x})$ happens to be degenerate, we can use the development given in Sec. 6 to test whether the kernel space of the approximating function intersects with the simplex. For the vast majority of the simplexes, the testing, and indeed probably the early part of the testing, will show a negative result and the search over the current simplex can be terminated immediately. The testing then goes to the next simplex in the array. If the testing does show the presence of the zero of $\mathbf{f}^L(\mathbf{x})$ in the simplex, this zero is then computed and taken to be the approximate location of a zero of $\mathbf{f}(\mathbf{x})$. This approximate zero can subsequently be improved upon by using the refining procedure to be discussed in Sec. 8. Implicit in the method is, of course, the assumption that the interval sizes are reasonably small so that the linear function determined by the function values at the vertices is an adequate approximation of $\mathbf{f}(\mathbf{x})$ over the simplex.

We now describe the first level search procedure itself.

(A) We first evaluate $\mathbf{f}_j = \mathbf{f}(\mathbf{x}_j), j = 0, 1, \ldots, n$.

(B) We next evaluate the basic $F$-determinants of (5.14), which have been abbreviated as $d^F_{(j)}$, by taking successively $j = 0, 1, \ldots, n$, and note whether $d^F_{(j)} < 0$, $> 0$, or $= 0$. After computing *each* of $d^F_{(j)}, j = 1, 2, \ldots, n$, we immediately compare its sign with the sign of the last previous nonvanishing basic $F$-determinant.

(B1) Any time the sign of the just computed $d^F_{(j)}$ is found to be opposite to that of the last nonvanishing determinant, one can immediately conclude that $\mathbf{x}^*$ of $\mathbf{f}^L(\mathbf{x})$ lies outside $s^n$ and terminate the testing.

(B2) If (B1) does not happen along the way, then at the end the $(n + 1)$ computed basic $F$-determinants will all be of the same sign with possibly some equal to zero. If none of them vanish, then the zero $\mathbf{x}^*$ of $\mathbf{f}^L(\mathbf{x})$ lies in $s^n$. If $d^F_{(k_1)} = d^F_{(k_2)} = \cdots = d^F_{(k_q)} = 0$, $1 \leqslant q \leqslant n$, then $\mathbf{x}^*$ of $\mathbf{f}^L(\mathbf{x})$ lies in a $(n - q)$-face of $s^n$ whose vertex index set is $\{k_1, k_2, \ldots, k_q\}^c$. If $q = n$, then $\mathbf{x}^*$ is at the vertex $\mathbf{x}_J$ where $J$ is associated with the only nonvanishing determinant $d^F_{(J)}$. Although there are indeed many possibilities within this case, they all say that the zero $\mathbf{x}^*$ of $\mathbf{f}^L(\mathbf{x})$ is in or on the boundary of the simplex. Therefore, we proceed to compute $\mathbf{x}^*$ by using (5.13) and (3.1). This is then taken to be the zeroth order approximation $\mathbf{x}^{(0)}$ to a zero of $\mathbf{f}(\mathbf{x})$. After this computation the testing on this simplex is terminated.

(B3) All $d^F_{(j)}, j = 0, 1, \ldots, n$, are zero. This means Rank $\mathbf{\Phi}^+(\mathbf{f}) < n + 1$ and Rank $\mathbf{\Phi}(\mathbf{f}) < n$. The case is degenerate; we proceed to (C) below.

*Remark.* The above procedure seems to ignore the case where $d^F_{(j)}$ may not be all of the same sign but their sum $[\det \mathbf{\Phi}^+(\mathbf{f})] = 0$. This also leads to a degenerate case. We ignore this case in the operational procedure because for this case Rank $\mathbf{\Phi}^+(\mathbf{f}) = n$ and Rank $\mathbf{\Phi}(\mathbf{f}) = n$ and, according to Theorem 6.1, there will be no zeros of $\mathbf{f}(\mathbf{x})$.

(C) We numerically operate on $\mathbf{\Phi}^+(\mathbf{f})$ by elementary operations to put it into a lower left triangular form with ones at the leading diagonal positions.

$$b_1 \begin{pmatrix} 1 & & & & & \\ \cdot & 1 & & & & \\ \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & 1 & & \\ \cdot & \cdot & \cdot & \cdot & \mathbf{0} \end{pmatrix} \tag{7.3}$$

Let the first $r_1$ columns be nonzero. Next, we numerically operate on $\mathbf{\Phi}(\mathbf{f})$ in a similar way resulting in

$$b\begin{pmatrix} 1 & & & & \\ \cdot & 1 & & & \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & 1 & \\ \cdot & \cdot & \cdot & \cdot & \mathbf{0} \end{pmatrix}. \tag{7.4}$$

Let the first $r$ columns of this matrix be nonzero.

(C1) $r_1 \neq r + 1$. In this case $\mathbf{f}^L(\mathbf{f})$ has no zeros.

(C2) $r_1 = r + 1$. In this case the smallest hyperplane in $F^n$ containing $V_F$ is of dimension $r$. The affine kernel subspace in $X^n$ is of dimension $n - r$. The first $r$ nonzero column vectors of (7.4) can be taken as a set of base vectors $\mathbf{c}_i$ in the subspace $\Lambda^r(V_F)$ and all the vectors $\mathbf{f}_i$, $i = 0, 1, \ldots, n$, can be expressed in terms of them, resulting in the coefficient set of $\alpha_{ij}$ of (6.8).

The next task is to test whether the affine kernel subspace $K^{n-r}$ intersects with $s^n$. When $r > 3$, to partition the convex set $\Gamma^r(V_F)$ into a number of nonoverlapping $r$-simplexes is not easy. A more viable scheme is simply to test all possible simplexes of $(r + 1)$ vertices from $V_F$. For each, say consisting of $\mathbf{f}_{k_0}, \mathbf{f}_{k_1}, \ldots, \mathbf{f}_{k_r}$, we first check to see whether

$$\det\begin{bmatrix} 1, & 1, \ldots, 1 \\ \boldsymbol{\alpha}_{k_0}, & \boldsymbol{\alpha}_{k_1}, \ldots, \boldsymbol{\alpha}_{k_r} \end{bmatrix} \tag{7.5}$$

is zero or not. If it is zero, this simplex is of dimension less than $r$ and it should be ignored. If it is not zero, then we carry out a test on this $r$-simplex to see whether the kernel subspace intersects with $s^n$ by following the procedure discussed in subsection (C) of Sec. 6.

When it is found that at a certain simplex $s^n$ in $X^n$, $\mathbf{f}^L(\mathbf{x})$ is degenerate and its kernel subspace intersects with $s^n$, then this fact is registered into a *roster of special simplexes*. This roster is to alert us that so far as the zeros of $\mathbf{f}(\mathbf{x})$ are concerned these locations in $X^n$ deserve further investigation.

In this manner we apply the first level search procedure to sweep the whole array of $n$-simplexes in $X^n$. When the sweep is completed, we would have obtained a *roster of discovered zeros* of $\mathbf{f}(\mathbf{x})$ for which we have their zeroth approximations and a *roster of special simplexes*.

**8. Refining the zeros.** The first level search procedure is seen to give us a very effective method to search through the whole domain of interest and to locate the zeros of given function. Generally, the zeros computed in this procedure may not be up to the required accuracy. Here, of course, once an approximate location of a zero has been found, any classical iteration methods can be used to improve the zeros. In this section we shall, however, describe a refining method which is based upon the same idea as the first level search procedure, namely: using a linear approximating function over an $n$-simplex. This gives the complete method an added element of unity.

The new refining method proceeds as follows. After having obtained an approximate location $\mathbf{x}^{(0)}$ of a zero from the first level search procedure we introduce a new set of interval sizes $h_i'$ such that

$$h_i' = \rho h_i, \qquad i = 1,2,\ldots,n, \quad 0 < \rho < 1, \tag{8.1}$$

where $\rho$ will be called the size reduction parameter. Next, we construct a new $n$-simplex with $\mathbf{x}^{(0)}$ at its barycentric [3]. The vertices of this $n$-simplex may be taken to be at

$$\begin{aligned}
\mathbf{x}_0 &= \mathbf{x}^{(0)} - \sum_{i=1}^{n} \frac{1}{i+1} h_i' \mathbf{e}_i, \\
\mathbf{x}_1 &= \mathbf{x}^{(0)} - \sum_{i=2}^{n} \frac{1}{i+1} h_i' \mathbf{e}_i + \frac{1}{2} h_1' \mathbf{e}_1, \\
&\vdots \\
\mathbf{x}_j &= \mathbf{x}^{(0)} - \sum_{i=j+1}^{n} \frac{1}{i+1} h_i' \mathbf{e}_i + \frac{j}{j+1} h_j' \mathbf{e}_j, \\
&\vdots \\
\mathbf{x}_n &= \mathbf{x}^{(0)} + \frac{n}{n+1} h_n' \mathbf{e}_n,
\end{aligned} \tag{8.2}$$

where $\mathbf{e}_i$, $i = 1,2,\ldots,n$, stands for $[0,\ldots,0,1,0,\ldots,0]^T$ with the only nonzero element at the $i$th position. It can be readily verified that

$$\mathbf{x}^{(0)} = \sum_{i=0}^{n} \frac{1}{n+1} \mathbf{x}_i \tag{8.3}$$

so that $\mathbf{x}^{(0)}$ is indeed the barycenter of this new smaller $n$-simplex. This simplex is entirely contained in an $n$-cube with sides $h_i'$, $i = 1,2,\ldots,n$. One can also show that the volume of this $n$-simplex is $(h_1' h_2' \cdots h_n')/n!$. Therefore, the reduction in volume when compared with the $n$-simplex used in the first level search procedure is by a factor $\rho^n$.

Having constructed this new $n$-simplex, we can evaluate the values of $\mathbf{f}(\mathbf{x})$ at the vertices of the new simplex. These function values determine an improved linear approximation function, because the new simplex is smaller in size. By using (5.13) and (3.1) we can then compute a new zero which will be denoted as $\mathbf{x}^{(1)}$, the first improved zero of the refining process. This completes the first iteration. As a check one could verify that the $(n+1)$ barycentric coordinates computed from (5.13) are all positive so that the new zero is inside the new simplex. This finding is a very reassuring thing to have but it is not a required part of this refining process. In any event, this will usually be the case unless the size reduction parameter has been taken too small.

With the improved zero $\mathbf{x}^{(1)}$ at hand we can repeat the refining process by constructing a new $n$-simplex, $\rho$ times smaller yet, with $\mathbf{x}^{(1)}$ at its barycenter. In this manner, we find by iteration a sequence of points $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots$, which are contained in a sequence of $n$-simplexes of decreasing size. After the $j$th iteration, the interval sizes will be $\rho^j h_i$, $i = 1,2,\ldots,n$. The iteration is terminated when $\rho^j h_i$ for all $i$ are less than the accuracies required by the zeros.

In the unlikely event that during the iteration process a linear approximating function is found to be degenerate, then the procedure of (C) of Sec. 7 should be followed. If the kernel subspace of that degenerate linear function intersects with the simplex, that simplex should be recorded in the roster of special simplexes.

As stated earlier, we describe this method of refining merely to show that the underlying idea for the first level search procedure could also be a basis for a refining process. When compared with the classical iteration procedures such as those based upon Newton's method, the present method requires, however, more computation for $n \geqslant 3$. Therefore, there is no justification to employ this method of refining except perhaps at the first few iteration steps.

**9. Concluding remarks.** We believe we have presented in this paper a useful method for locating the simple zeros of a vector function. The method complements very naturally the various conventional methods of finding zeros which require a prior knowledge of the approximate locations of the zeros. Experience with the method up to now shows the method to be highly effective and efficient. The discussions given in this paper are all on the basic aspects of the method. Many variations of the basic scheme are possible and they will further increase the efficacy of the method.

Finally, we wish to point out in passing that the discussion given in Sec. 5 is intimately linked to the index theory of vector fields, [5, 6]. In this connection we might also mention a recent work [7] in which an index theory for cell functions is presented. Although specific results from [7] are not used in this paper, they did influence strongly the development of the basic idea behind the present proposed method.

**Acknowledgments.** This material is based upon work supported by the National Science Foundation under Grant No. MEA-8019274.

## References

[1] P. S. Aleksandrov, *Combinatorial topology*, Vol. 1, Graylock Press, Rochester, N.Y., 1956
[2] K. M. Brown, *A quadratically convergent Newton-like method based upon Gaussian elimination*, SIAM J. Numer. Anal. **6**, 560–569 (1969)
[3] S. S. Cairns, *Introductory topology*, Ronald Press, New York, 1968
[4] B. Carnahanm H. A. Luther and H. O. Wilkes, *Applied numerical methods*, Wiley, New York, 1969
[5] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill, New York, 1955
[6] C. S. Hsu, *A theory of index for dynamical systems of order higher than two*, J. Appl. Mech., **47**, 421–427 (1980)
[7] C. S. Hsu and W. H. Leung, *Singular entities and an index theory for cell functions*, J. Math. Anal. Appl. (to appear)
[8] S. Lefschetz, *Introduction to topology*, Princeton University Press, Princeton, N.J., 1949
[9] K. Nomizu, *Fundamentals of linear algebra*, McGraw-Hill, New York, 1966
[10] S. Perlis, *Theory of matrices*, Addison-Wesley, Reading, Mass., 1952