NASA-CR-194502

IN-32-CR
186566
84 P

JPL Publication 91-45

Using a Focal-Plane Array to Estimate Antenna Pointing Errors

S. Zohar
V. A. Vilnrotter

October 25, 1991

NASA

1. Report No. 91-45	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Using a Focal-Plane Array to Estimate Antenna Pointing Errors		5. Report Date October 25, 1991
		6. Performing Organization Code
7. Author(s) S. Zohar, V. A. Vilnrotter		8. Performing Organization Report No.
9. Performing Organization Name and Address JET PROPULSION LABORATORY California Institute of Technology 4800 Oak Grove Drive Pasadena, California 91109		10. Work Unit No.
		11. Contract or Grant No. NAS7-918
		13. Type of Report and Period Covered JPL Publication
12. Sponsoring Agency Name and Address NATIONAL AERONAUTICS AND SPACE ADMINISTRATION Washington, D.C. 20546		
		14. Sponsoring Agency Code RE210 BG-310-30-70-89-01

15. Supplementary Notes

16. Abstract

We consider here the use of extra collecting horns in the focal plane of an antenna as a means of determining the Direction of Arrival (DOA) of the signal impinging on it, provided it is within the antenna beam. Our analysis yields a relatively simple algorithm to extract the DOA from the horns' outputs. We also develop here an algorithm which, in effect, measures the thermal noise of the horns' signals and determines its effect on the uncertainty of the extracted DOA parameters.

Both algorithms have been implemented in software and tested in simulated data. Based on these tests, we conclude that this is a viable approach to the DOA determination.

Though the results obtained here are of general applicability, the particular motivation for the present work is their application to the pointing of a mechanically deformed antenna. It is anticipated that the pointing algorithm developed here for a deformed antenna could be obtained as a small perturbation of the algorithm developed here for an undeformed antenna. In this context, it should be pointed out that, with a deformed antenna, the array of horns and its associated circuitry constitute the main part of the deformation-compensation system. In this case, the pointing system proposed here may be viewed as an additional task carried out by the deformation-compensation hardware.

17. Key Words (Selected by Author(s))	18. Distribution Statement Unclassified; unlimited

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 83	22. Price

JPL 0184 R 9/83

I

Using a Focal-Plane Array to Estimate Antenna Pointing Errors

S. Zohar
V. A. Vilnrotter

October 25, 1991

NASA

ABSTRACT

We consider here the use of extra collecting horns in the focal plane of an antenna as a means of determining the Direction Of Arrival (DOA) of the signal impinging on it, provided it is within the antenna beam. Our analysis yields a relatively simple algorithm to extract the DOA from the horns' outputs. We also develop here an algorithm which, in effect, measures the thermal noise of the horns' signals and determines its effect on the uncertainty of the extracted DOA parameters.

Both algorithms have been implemented in software and tested on simulated data. Based on these tests, we conclude that this is a viable approach to the DOA determination.

Though the results obtained here are of general applicability, the particular motivation for the present work is their application to the pointing of a mechanically deformed antenna. It is anticipated that the pointing algorithm for a deformed antenna could be obtained as a small perturbation of the algorithm developed here for an undeformed antenna. In this context, it should be pointed out that, with a deformed antenna, the array of horns and its associated circuitry constitute the main part of the deformation-compensation system. In this case, the pointing system proposed here may be viewed as an additional task carried out by the deformation-compensation hardware.

CONTENTS

CONTENTS (Continued)

I.   INTRODUCTION

Consider a circular-paraboloid reflector antenna in its reception mode. We have here a large "dish" which collects the energy arriving from a distant source and concentrates it (directly or via a Cassegrain system) onto the single collecting horn, which provides the antenna output. Conceptually, one can distinguish two basic planes in this system: the aperture plane, where the field has the uniform distribution of a plane wave, and the focal plane, where the field distribution is quite different, being characterized by a very pronounced peak at the location of the collecting horn.

With a properly designed antenna, most of the energy incident on the antenna aperture is concentrated into this peak and thus fed to the output via the collecting horn. When such an antenna is subjected to a mechanical deformation, the field distribution in the focal plane will be diffuse so that the collecting horn will collect only part of the total energy incident on the antenna and we observe a signal loss.

While the source of the mechanical deformation is of no concern to us here, we should point out a practically important scenario where such deformation has to be dealt with. We are referring to the situation where we try to extend upward the frequency coverage of an already-built functional antenna. In this case, the antenna whose mechanical deviation from the ideal is acceptable within its design frequency range will now be labeled "mechanically deformed" at the shorter wavelengths.

One possible way of compensating for at least part of the deformation signal loss is to augment the single collecting horn with an array of additional horns closely packed around it and produce the overall system output as a weighted sum of all the horn outputs [1]. The main challenge in such a design is to determine the optimal combining weights and implement them in real time.

When such a system is properly pointed to extract the maximum power from a desired source, it may be considered equivalent to a standard undeformed antenna of a somewhat smaller aperture, and one tends to presume that the deformation and the array of horns compensating for it are totally transparent to the user. This, however, is not the case since the behavior of the system when it is turned slightly away from this optimal direction is quite different from that of a standard antenna. For example, the mechanical deformation might be such that maximal output is produced when the antenna is pointed slightly off the actual direction of the source. This, then, raises a basic question: Given the direction of the desired source, how will such an antenna system be pointed to extract the maximal power from it?

Obviously, the availability of the outputs of the individual horns of the collecting array provides the basis for solving this problem. We propose here a two-phase approach: In phase 1 we assume that the deformation is negligible and develop an algorithm which processes the outputs of the individual horns to yield the Direction Of Arrival (DOA) of a plane wave impinging on the antenna.

In phase 2 we consider the case of the actual deformed antenna as a small perturbation of the undeformed case and look for the modifications and additional parameters that have to be provided so that the DOA can still be extracted. In this report, we cover phase 1.

Throughout this report we use the acronym DOA to represent the entity we want to extract, namely, the Direction Of Arrival of the signal. Strictly speaking, this is a precise characterization. However, there is a large amount of literature in which "extracting the DOA" has the additional implication of starting with a blank page, that is, we do not have even a coarse a priori estimate of the DOA. We want to stress at the outset that this is not the usage intended here. The algorithm developed in this report is applicable only when the antenna is already pointed at the vicinity of the source. Thus, this algorithm handles the task of fine pointing after coarse pointing has already been effected by other means.

The general outline of the report is as follows: In Section II we develop the basic strategy underlying the DOA extraction. We show there that the DOA is determined by the gradient of the phase angle of the field entities across the antenna aperture. Since the aperture-plane entities are not directly available to us, we have to infer them from the focal-plane "measurements," namely, the outputs of the focal-plane horns. This is tackled in Section III.

From this point on, our analysis is strongly influenced by the fact that the signals available to us are noisy. To (partially) overcome the effect of this noise, we have to increase the number of measurements and are thus naturally led to a least-squares formulation of the problem (number of measurements exceeds number of unknowns). It turns out that, in the least-squares context, it is more convenient to regard the DOA extraction as the determination of the orientation of a wavefront of the incident wave. The linear operator underlying this formulation is derived in Section IV. In Section V we study the structure of this linear operator using the powerful concepts of Singular Value Decomposition (SVD). We find that the representation of this operator yielded by our formulation is so simple that it allows us to obtain a closed-form solution of the least-squares problem (no matrix inversion; no SVD subroutines).

At this point of the analysis we are in possession of an algorithm to obtain the best (least-squares) estimate of the DOA. The next three sections are devoted to the much more challenging task of determining the reliability of this estimate. We express this reliability as follows: The estimated DOA is representable as a point ($\hat{D}$) on the unit sphere. We look for the shape and size of the region around this point in which the point D representing the true DOA is guaranteed to fall with a prescribed probability.

The first step towards this goal is the determination of the covariance matrix of the least-squares solution vector. In Section VI we express this covariance matrix in terms of the covariance matrix of the parameters at the input of the least-squares problem. Here we encounter a

2

major challenge: The input parameters of the least-squares problem are the phases of the aperture-plane field entities. But the noise is injected in the focal plane and the transformation between these two planes transforms the diagonal noise covariance matrix of the focal-plane into a nondiagonal noise covariance matrix in the aperture plane where the least-squares processing is applied.

Section VII is devoted to the determination of this nondiagonal noise covariance matrix. The mathematical manipulations in this section are quite complex, but we have provided here a sufficiently detailed step-by-step derivation to compensate for that.

In Section VIII we handle the extraction of the confidence region from the (now available) output covariance matrix. The final result here is an algorithm to determine the confidence region for any probability.

The two algorithms developed in this report have been implemented in software and incorporated in a very flexible program which allows their testing with simulated or real data, with or without added noise, and with the easy selection of a large number of operational parameters. Preliminary results with simulated data indicate that these algorithms constitute a technically sound approach to the DOA determination. Some of these results are discussed in Section IX.

In Section X we present a brief overview and a discussion of some worthwhile extensions of this work.

Throughout this report, we use the formalism of dyadics to handle linear operators. When combined with the concepts of SVD, this provides a very powerful tool which has proved to be quite useful in mathematical manipulations as well as on the conceptual level. For those not familiar with this formalism, we provide a short but sufficient summary of the subject in Appendix A.

Some of the detailed arguments in Section III require knowledge of the attenuation of the focal-plane field as a function of distance from the optical axis when the incident plane wave is off boresight. In Appendix B we derive this result. An incidental by-product of this derivation is the possibility that it may provide the basis for a DOA algorithm which is quite different from the one developed in this report. This is briefly discussed in this appendix.

The number of variables in this work is quite large. To reduce the number of required symbols, we have made extensive use of the caret (^) to allow use of the same letter for two different but related entities. Even so, we find that some symbols have to be assigned different meanings in the different sections. The quite different contexts, however, should rule out any ambiguity.

3

II. THE BASIC STRATEGY

The clue to the DOA extraction is the very special variation of the field entities of a plane wave as functions of time and space. Specifically, for a plane wave of radial frequency ω (and corresponding wavelength λ), all components of the electric and magnetic fields vary as

$$e^{i(\bar{k}\cdot\bar{r}-\omega t)} \tag{2.1}$$

where $\bar{r}$ is the position vector of the point where the field is specified (throughout this report, we use overbars to denote vectors), t is the relevant time, and the vector $\bar{k}$ is the propagation vector, namely, a vector of magnitude $k = 2\pi/\lambda$ pointing in the direction of propagation of the plane wave. (Note: we adopt here the convention that the symbol of a vector entity without the overbar stands for the magnitude of that vector).

Note that the dimensionless entity

$$\alpha \equiv \bar{k} \cdot \bar{r} - \omega t \tag{2.2}$$

appearing in (2.1) is a phase angle. Thus, if we have sensors which continuously monitor a certain component of the field in various locations in space, then all of them will display a sinusoidal variation with time with the same radial frequency but with different phases, the phase of each sensor being given by the α corresponding to its position vector $\bar{r}$ (see (2.2)). The DOA we are after will be specified in terms of a unit vector $(\bar{n})$ pointing at the source of the wave. From the definition of $\bar{k}$, it is obvious that

$$\bar{k} = -(2\pi/\lambda)\,\bar{n} \tag{2.3}$$

and our problem is essentially that of determining $\bar{k}$. With that in mind, we return to (2.2) and consider the gradient of the phase angle α, namely,

$$\bar{\nabla}\alpha = \bar{k} \cdot \bar{\nabla}\bar{r} \tag{2.4}$$

But

$$\bar{\nabla}\bar{r} = \bar{\bar{I}} \tag{2.5}$$

where $\overline{\overline{I}}$ is the identity dyadic. (Throughout this report, we use double overbars to denote dyadics. See Appendix A for a short but sufficient summary of the subject of dyadics as well as a simple derivation of (2.5).) Hence,

$$\bar{k} = \overline{\nabla}\alpha \qquad (2.6)$$

and, using (2.3),

$$\bar{n} = -\frac{\lambda}{2\pi} \overline{\nabla}\alpha \qquad (2.7)$$

We conclude that measuring the phase differences across an array of sensors monitoring the plane wave will yield the desired DOA. This, then, specifies the main points of our basic strategy:

(1) We consider the focal-plane horns as sensors monitoring the focal-plane field.

(2) We use the Fourier-transform relationship between the aperture-plane field and the focal-plane field to obtain from the horns' outputs a field distribution across the aperture plane.

(3) Applying (2.7) to the phases of the aperture-plane distribution, we obtain the desired DOA.

In principle, (2.7) requires a three-dimensional phase distribution rather than the planar distribution obtained in the scheme outlined above. However, in our case, a planar distribution is sufficient, as we proceed to show now.

Let us introduce an antenna-fixed Cartesian frame $\{\bar{h}_j\}_{j=1}^3$ as shown in Fig. 1. In this frame, all aperture-plane points are characterized by zero $\bar{h}_3$ components. This motivates us to split $\overline{\nabla}\alpha$ as follows:

$$\overline{\nabla}\alpha = (\bar{h}_3\bar{h}_3) \cdot (\overline{\nabla}\alpha) + (\overline{\overline{I}} - \bar{h}_3\bar{h}_3) \cdot (\overline{\nabla}\alpha) \qquad (2.8)$$

Let us denote

$$\overline{\overline{I}} - \bar{h}_3\bar{h}_3 \equiv \overline{\overline{P}}_3 \qquad (2.9)$$

Y

Z

X-AXIS POINTING INTO PAGE

$\bar{h}_1$ POINTING INTO PAGE

$\bar{h}_2$

$\bar{h}_3$
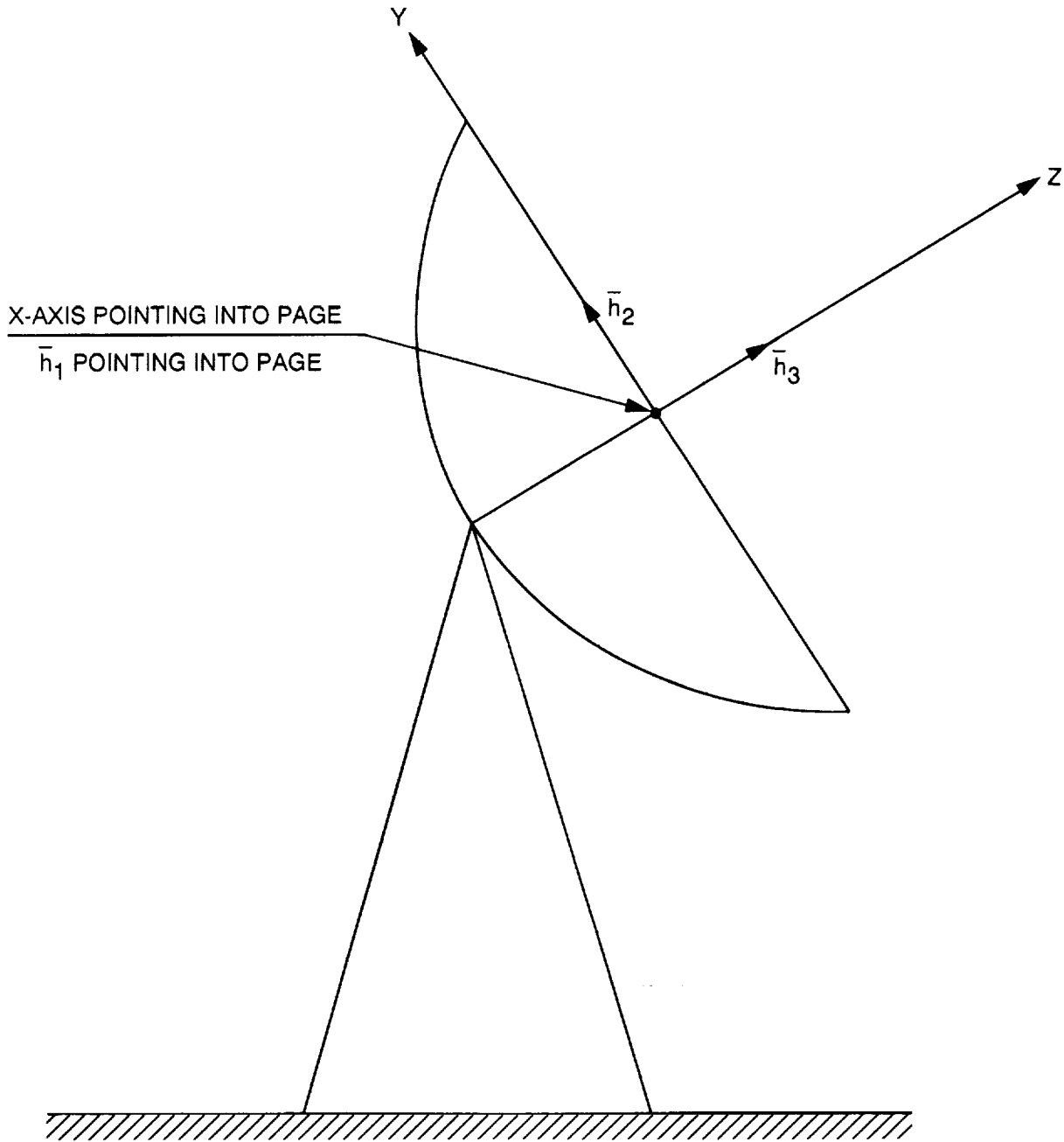
Fig. 1. The antenna-fixed frame $\{\bar{h}_i\}_{i=1}^3$.

$\bar{\bar{P}}_3$ is a projection operator which extracts from any vector its part orthogonal to $\bar{h}_3$. Thus,

$$\bar{\nabla}\alpha = (\bar{h}_3\bar{h}_3) \cdot (\bar{\nabla}\alpha) + \bar{\bar{P}}_3 \cdot (\bar{\nabla}\alpha) \qquad (2.10)$$

and the entity available to us from the aperture-plane distribution is not $\bar{\nabla}\alpha$ but

$$\bar{\bar{P}}_3 \cdot (\bar{\nabla}\alpha) \equiv \bar{\nabla}_S \alpha \qquad (2.11)$$

Such an entity is usually referred to as a "surface gradient."

To see the effect of this shortcoming, let us apply $\bar{\bar{P}}_3$ to both sides of (2.7):

$$\bar{\bar{P}}_3 \cdot \bar{n} = -\frac{\lambda}{2\pi} \bar{\nabla}_S \alpha \qquad (2.12)$$

Now we express $\bar{n}$ in terms of the $\bar{h}_i$ frame

$$\bar{n} = \sum_{i=1}^{3} n_i \bar{h}_i \qquad (2.13)$$

Hence,

$$\bar{\bar{P}}_3 \cdot \bar{n} = n_1 \bar{h}_1 + n_2 \bar{h}_2 \equiv \bar{n}_S \qquad (2.14)$$

and the final result is

$$\bar{n}_S = -\frac{\lambda}{2\pi} \bar{\nabla}_S \alpha \qquad (2.15)$$

Note that

$$\bar{n} = \bar{n}_S + n_3 \bar{h}_3 \qquad (2.16)$$

7

with

$$n_3 = \pm \sqrt{1 - n_s^2} \qquad (2.17)$$

Thus, once we have computed $\bar{n}_s$, the location of the source is constrained to just two mathematically possible alternatives, one in front of the antenna $(n_3 > 0)$ and one behind it $(n_3 < 0)$. Since the choice between these two is obvious, we conclude that the surface gradient is sufficient to determine the DOA in our case.

In the next section, we take up step (2) of our basic strategy, namely, determining the unavailable aperture-plane distribution from the available focal-plane distribution.

III. EXTRACTING APERTURE-PLANE FIELD PARAMETERS FROM FOCAL-PLANE MEASUREMENTS

Consider the field distributions in the two basic planes associated with our antenna, namely, the aperture plane and the focal plane. In principle, we can obtain one distribution from the other through the use of a two-dimensional Fourier transformation. This is based on the Fresnel approximation and will be phrased more precisely later on. First, however, we have to deal with some preliminaries concerning two-dimensional Fourier transforms in general.

We start with a function of time $\hat{g}(t)$ and its Fourier transform $g(f)$ given by

$$g(f) = \int_{-\infty}^{\infty} \hat{g}(t) e^{-i2\pi ft} dt \tag{3.1}$$

Notice that the obvious requirement that the argument of the exponential function be dimensionless is satisfied by the fact that the dimension of frequency is 1/time.

Now let us change the scenario by assuming that the function $\hat{g}$ represents a distribution of some entity along a line in space so that $\hat{g}(t)$ is replaced now by a function of $\hat{x}$ where $\hat{x}$ specifies position along that line. When we want to determine the Fourier transform of this distribution, we are faced with the choice of the substitution to be adopted for the frequency variable of (3.1). Goodman [2] (and many others) chooses what is referred to as "spatial frequency" measured in units of 1/length. While this satisfies the above-mentioned dimensionality constraint, it is conceptually very unsatisfactory because usually both $\hat{g}$ and g are functions of position (i.e., $\hat{g}$ is a function of the position variable $\hat{x}$, while g is a function of the position variable x) .

We adopt here a different formulation: We preserve the dimensionality constraint by introducing a reference length (L) and adopting the dimensionless entity (length $/L$) to be the argument of both $\hat{g}$ and g. Specifically, we let

$$\hat{X} = \hat{x}/L \tag{3.2}$$

$$X = x/L \tag{3.3}$$

and the Fourier transform is now expressed as

$$g(X) = \int_{-\infty}^{\infty} \hat{g}(\hat{X}) e^{-i2\pi X\hat{X}} d\hat{X} \tag{3.4}$$

Generalizing to the two-dimensional case, we obtain

$$g(\bar{R}) = \iint_{-\infty}^{\infty} \hat{g}(\hat{\bar{R}}) e^{-i2\pi\bar{R}\cdot\hat{\bar{R}}} \, d\hat{S} \tag{3.5}$$

where $d\hat{S}$ is a (normalized) area element in the input plane spanned by $\hat{\bar{R}}$. In words, the spatial distribution $g(\bar{R})$ is the two-dimensional Fourier transform of the spatial distribution $\hat{g}(\hat{\bar{R}})$ where

$$\hat{\bar{R}} = \hat{\bar{r}}/L \tag{3.6}$$

$$\bar{R} = \bar{r}/L \tag{3.7}$$

It should be pointed out that the reference length L is not an arbitrary parameter. The physics (or other underlying theory) which yields the Fourier transformation property associating the functions g and $\hat{g}$ will also yield a specific value for the reference length L. The thin-lens image transformation which concerns us here is a case in point. From Goodman's formulation of this transformation (equation (5-15) of [2]), we infer that for an antenna of focal length F, operating at the wavelength λ, the reference length L is simply the geometric mean of these two lengths, that is,

$$L = \sqrt{\lambda F} \tag{3.8}$$

We turn now to the precise formulation of the thin-lens image transformation which we plan to apply to our DOA problem. Let

$$\hat{\bar{r}} = \text{position vector in the aperture plane } (\hat{\bar{r}} = \bar{0} \text{ on optical axis}) \tag{3.9}$$

$$\bar{r} = \text{position vector in the focal plane } (\bar{r} = \bar{0} \text{ on optical axis}) \tag{3.10}$$

$$\hat{u} = \text{a selected field component in the aperture plane} \tag{3.11}$$

$$u = \text{the corresponding field component in the focal plane} \tag{3.12}$$

Then, the focal-plane field distribution $u(\bar{R})$ is obtainable from the aperture-plane distribution $\hat{u}(\hat{\bar{R}})$ as follows:

$$u(\bar{R})\ e^{-i\pi R^2} = e^{i\alpha} \iint\limits_{-\infty}^{\infty} \hat{u}(\hat{\bar{R}})\ e^{-i2\pi \bar{R}\cdot\hat{\bar{R}}}\ d\hat{S} \qquad (3.13)$$

where α is a constant phase. This formulation is an exact equivalent of equation (5-15) of [2]. Comparing (3.13) to (3.5), we note that, apart from the trivial constant phase shift α, $\left[u(\bar{R})\ e^{-i\pi R^2}\right]$ is the Fourier transform of $\hat{u}(\hat{\bar{R}})$. Equivalently, we may say that $u(\bar{R})$ is approximated by the Fourier transform of $\hat{u}(\hat{\bar{R}})$ with the factor $e^{i(\pi R^2 + \alpha)}$ accounting for the deviation from an exact Fourier transform.

In our application, we start with a focal-plane distribution and look for the aperture-plane distribution that caused it.

Hence we invert (3.13), getting

$$\hat{u}(\hat{\bar{R}}) = e^{-i\alpha} \iint\limits_{-\infty}^{\infty} \left[u(\bar{R})\ e^{-i\pi R^2}\right]\ e^{i2\pi \hat{\bar{R}}\cdot\bar{R}}\ dS \qquad (3.14)$$

This formulation provides the basis for our algorithm. In its actual implementation, however, we have to replace the continuous Fourier transform by a (two-dimensional) Discrete Fourier Transform (DFT). We proceed now to set the stage for this modification. Our starting point is the focal-plane array of closely packed circular horns shown in Fig. 2. Let $\bar{R}_j$ be the normalized position vector of the center of the aperture of horn j and let $v(\bar{R}_j)$ be its (complex) voltage output. We consider the selected field entity at point $\bar{R}_j$ to be proportional to $v(\bar{R}_j)$.

· In Fig. 2, we show three concentric "rings" of horns surrounding the central horn at the origin. Though our initial actual implementation involves only one ring, we formulate our analysis in terms of an arbitrary number of rings (N). In this general case, the total number of horns in the array (J) is given by

$$J = 1 + 3N(N + 1) \qquad (3.15)$$

Thus, we have a set of J focal-plane voltages $v(\bar{R}_j)$, which we want to process in order to get a set of aperture-plane "voltages" $\hat{v}(\hat{\bar{R}}_k)$. In this context (and ignoring irrelevant scale factors), (3.14) transforms to
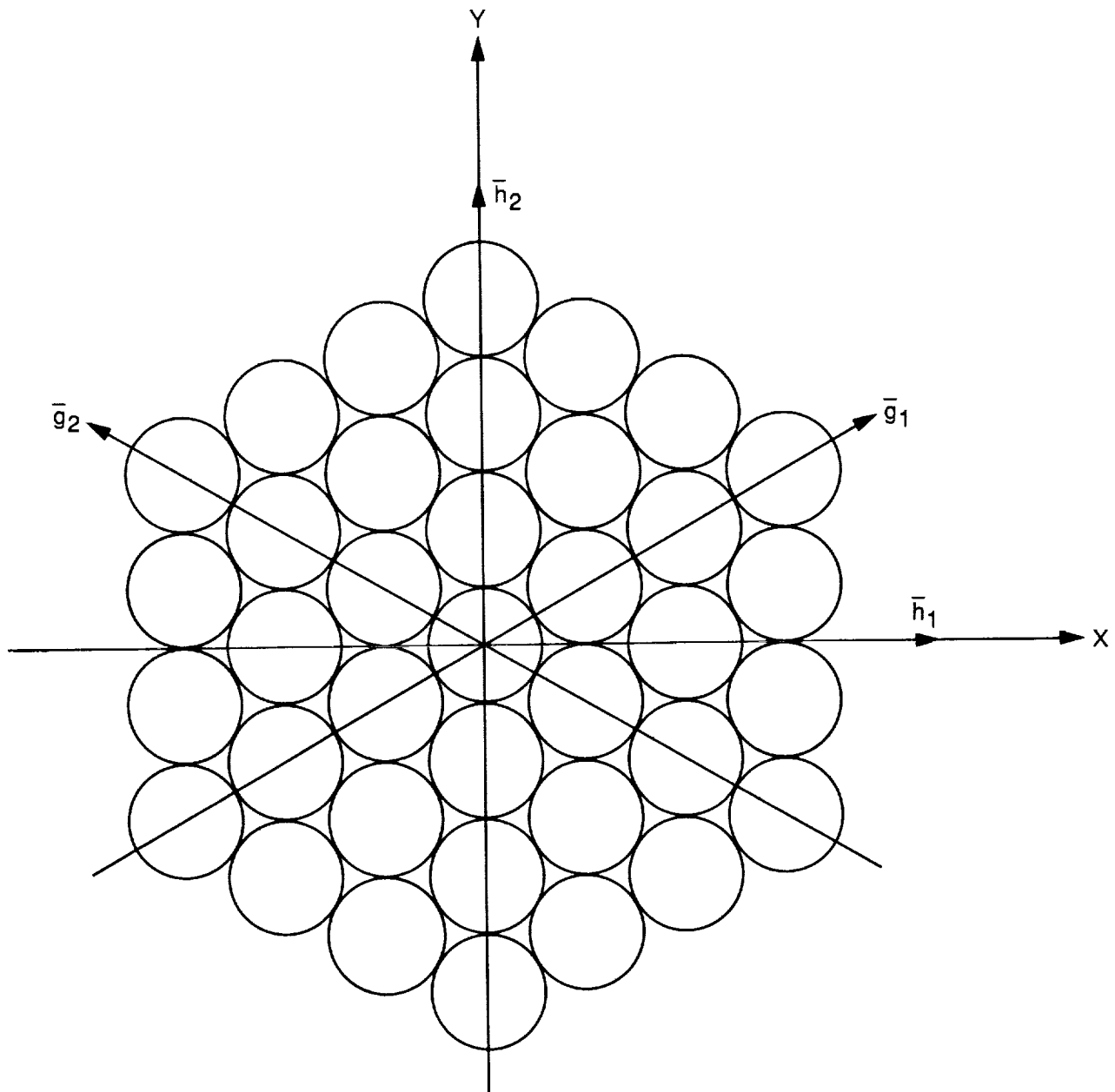
Fig. 2. The horns' configuration in the focal plane.

$$\hat{v}(\hat{\bar{R}}_k) = \sum_{j=1}^{J} \left[v(\bar{R}_j) \, e^{-i\pi R_j^2} \right] e^{i2\pi\hat{\bar{R}}_k \cdot \bar{R}_j} \qquad (3.16)$$

Equivalently,

$$\hat{v}(\hat{\bar{R}}_k) = \sum_{j=1}^{J} e^{i\pi\bar{R}_j \cdot (2\hat{\bar{R}}_k - \bar{R}_j)} \, v(\bar{R}_j) \qquad (3.17)$$

As in the case of the one-dimensional DFT, there are two important aspects of the transition from the continuous (3.14) to the discrete (3.16). First, the fact that we use discrete samples of the input function rather than the function itself means that the computed transform is periodic and could (potentially) be corrupted by aliasing. However, whereas in the one-dimensional case we specify this periodicity by a single parameter, say, the period or the frequency, in the two-dimensional case we have to be concerned with the size and shape of the repeating cell.

The second aspect derives from the finite summation in (3.16), that is, the fact that we have only a finite number of sampling points (horns) in the focal plane. Usually, this is handled by the introduction of a window function. However, as shown in Appendix B, for our adopted parameters, the terms excluded from (3.16) are very small and may be simply ignored. Thus, the only issue to be addressed in the implementation of the discrete (3.16) is the effect of the aperture-plane periodicity imposed by it.

In preparation for this task, we introduce now some required notational conventions. The focal-plane distribution of horns suggests that the nonorthogonal $\bar{g}_i$ frame shown in Fig. 2 is more suitable than the more conventional orthonormal $\bar{h}_i$ frame shown there. If we denote the horn diameter by d, then any $\bar{R}_j$ is expressible in terms of components in the $\bar{g}_i$ frame as follows:

$$\bar{R}_j = \bar{R}_{j_1, j_2} = (d/L) \, (j_1 \bar{g}_1 + j_2 \bar{g}_2) \qquad (3.18)$$

Note that j_1, j_2 are integers ranging over the interval $(-N, N)$ (N is the number of rings), though not all such positions are occupied. For example, while there is a horn at position $\bar{R}_{3,1}$ there is no horn at position $\bar{R}_{3,-1}$ (see Fig. 2).

As indicated in (3.18), we find it convenient to designate one and the same position vector as either a singly indexed entity or a doubly indexed entity. This, of course, implies a unique

correspondence between j and the pair (j_1, j_2). Indeed, in the software implementation of the algorithms developed here, we specifically establish such a relationship. However, in the theoretical manipulations in this report, it is enough to know that such a relationship exists without having to specify its particular form.

It is well known that manipulations involving nonorthogonal frames (such as $\{\bar{g}_i\}_{i=1}^2$) are appreciably simplified by the introduction of a reciprocal (dual) set of base vectors $\{\bar{g}^i\}$ which satisfy the cross-orthonormality condition

$$\bar{g}_i \cdot \bar{g}^j = \delta_{ij} \qquad (i, j = 1, 2) \tag{3.19}$$

Both sets of base vectors for our case are shown in Fig. 3. Note that while any vector may be represented in either frame, a judicious selection of frames will yield simpler expressions. In particular, we choose to represent an arbitrary point $\hat{\bar{R}}$ in the aperture plane in terms of the $\bar{g}^i$ frame as follows:

$$\hat{\bar{R}} = (L/d)(\rho_1 \bar{g}^1 + \rho_2 \bar{g}^2) \tag{3.20}$$

where ρ_1, ρ_2 are arbitrary scalars.

We exploit this representation now to establish the aperture-plane periodicity mentioned above. From (3.16) we see that the dependence of $\hat{v}(\hat{\bar{R}})$ on $\hat{\bar{R}}$ is effected through the dot product $\hat{\bar{R}} \cdot \bar{R}_j$, which we evaluate using the representations (3.18), (3.20) and applying (3.19)

$$\hat{\bar{R}} \cdot \bar{R}_j = \rho_1 j_1 + \rho_2 j_2 \tag{3.21}$$

Now we look for a $(\Delta \hat{\bar{R}})$ such that

$$\hat{v}(\hat{\bar{R}} + \Delta \hat{\bar{R}}) = \hat{v}(\hat{\bar{R}}) \tag{3.22}$$

From (3.21), (3.20), and (3.16), we infer that

$$\Delta \hat{\bar{R}} = k_1 \bar{p}^1 + k_2 \bar{p}^2 \tag{3.23}$$

where k_1, k_2 are arbitrary integers and

$$\bar{p}^i = (L/d) \bar{g}^i \qquad (i = 1, 2) \tag{3.24}$$

14

$$\overline{g}_1 = \frac{\sqrt{3}}{2}\,\overline{h}_1 + \frac{1}{2}\,\overline{h}_2$$

$$\overline{g}_2 = -\frac{\sqrt{3}}{2}\,\overline{h}_1 + \frac{1}{2}\,\overline{h}_2$$

$$\overline{g}^1 = \frac{1}{\sqrt{3}}\,\overline{h}_1 + \overline{h}_2$$

$$\overline{g}^2 = -\frac{1}{\sqrt{3}}\,\overline{h}_1 + \overline{h}_2$$

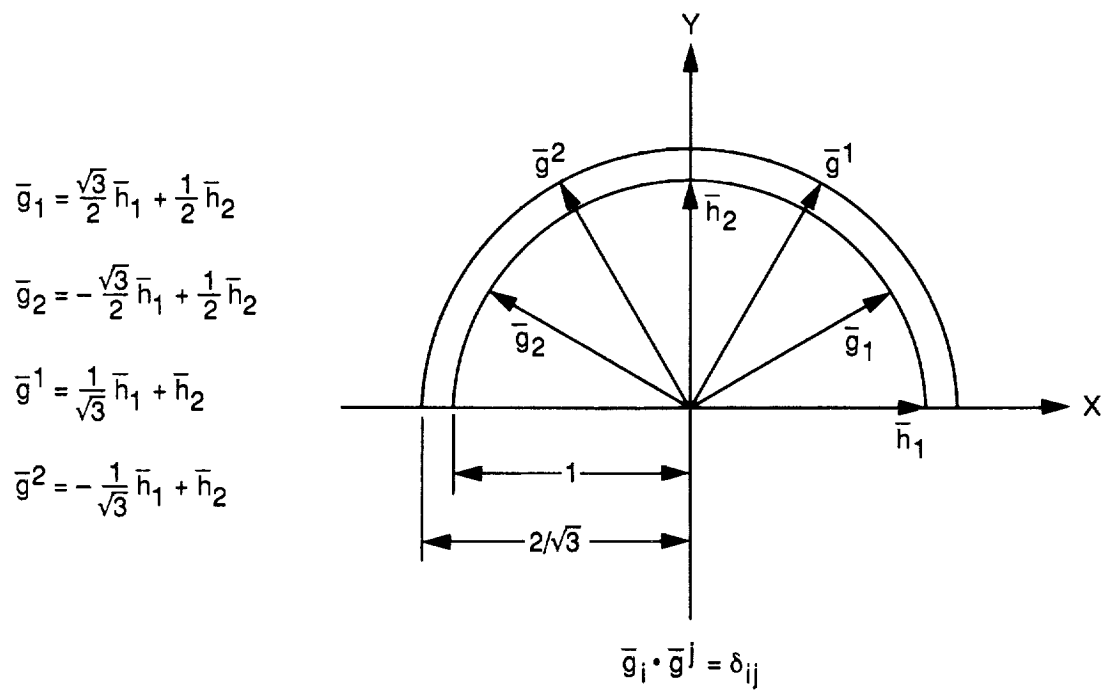$$\overline{g}_i \cdot \overline{g}^j = \delta_{ij}$$

Fig. 3.   The reciprocal (dual) vector sets.

15

Recalling now (see Fig. 3) that

$$g^i = 2/\sqrt{3} \qquad (i = 1, 2) \qquad (3.25)$$

we get for the length of $\bar{p}^i$

$$p^i = (2/\sqrt{3})(L/d) \qquad (i = 1, 2) \qquad (3.26)$$

This establishes the basic periodic structure which is shown in Fig. 4. We see here that the aperture plane is tiled with 60° rhombi with sides of (normalized) length $(2/\sqrt{3})(L/d)$. The established periodicity means that, using (3.16) to compute $\hat{v}$, we will find that the distribution of $\hat{v}$ over a rhombic tile is the same for all tiles.

Now we use this information to guide us in the selection of the set of points $\hat{\bar{R}}_k$ where the aperture-plane voltages are to be computed. Obviously, all of the $\hat{\bar{R}}_k$ points should be constrained to a single rhombic tile. Furthermore, the approximation error of the Fresnel approximation (equation (4-10) of [2]), which forms the basis of (3.13), increases with the distance from the optical axis. Hence, the selected rhombic tile should be the one centered at the origin.

Recall now that, in the focal plane, we have placed $(2N + 1)$ points equally spaced along each $\bar{g}_i$ axis (Fig. 2). Generalizing from common practice in the application of the one-dimensional DFT, one could argue that a reasonable strategy is to select the $\hat{\bar{R}}_k$ aperture-plane points to fully cover the origin rhombus with $(2N + 1)$ points equally spaced along each $\bar{g}^i$ axis as shown (for $N = 3$) in Fig. 5, for a total number of points (K) given by

$$K = (2N + 1)^2 \qquad (3.27)$$

Note that, unlike Fig. 2 where the circles represent the circular apertures of the conical collecting horns, the circles here are just convenient representation artifacts to highlight the fixed spacing of the computed aperture-plane points: there is an $\hat{\bar{R}}_k$ point at the center of each circle.

Our main concern in the selection of the aperture-plane points is the avoidance of the detrimental effects of aliasing. As we shall see shortly, in some practical cases, this dictates a higher density of aperture-plane points. In view of that, we deviate from the configuration of Fig. 5 by the introduction of the extra scaling parameter s. Specifically, we propose to pack the K points inside a smaller, central, 60° rhombus given by

$$\text{(the normalized side of the } \hat{\bar{R}}_k \text{ bounding rhombus)} = s \left(\frac{2}{\sqrt{3}}\right) \left(\frac{L}{d}\right) \qquad (3.28)$$

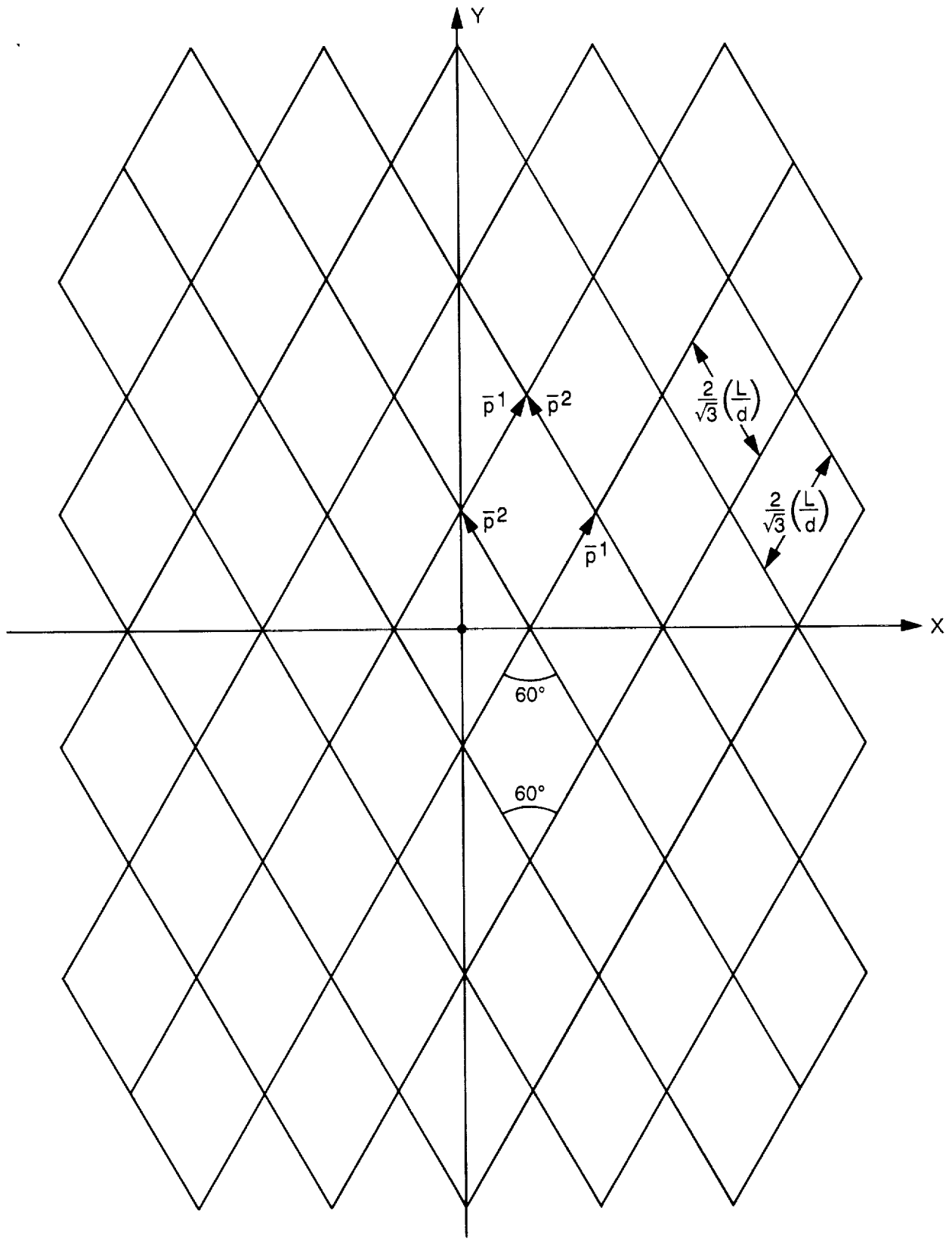(see Fig. 6) and impose the condition

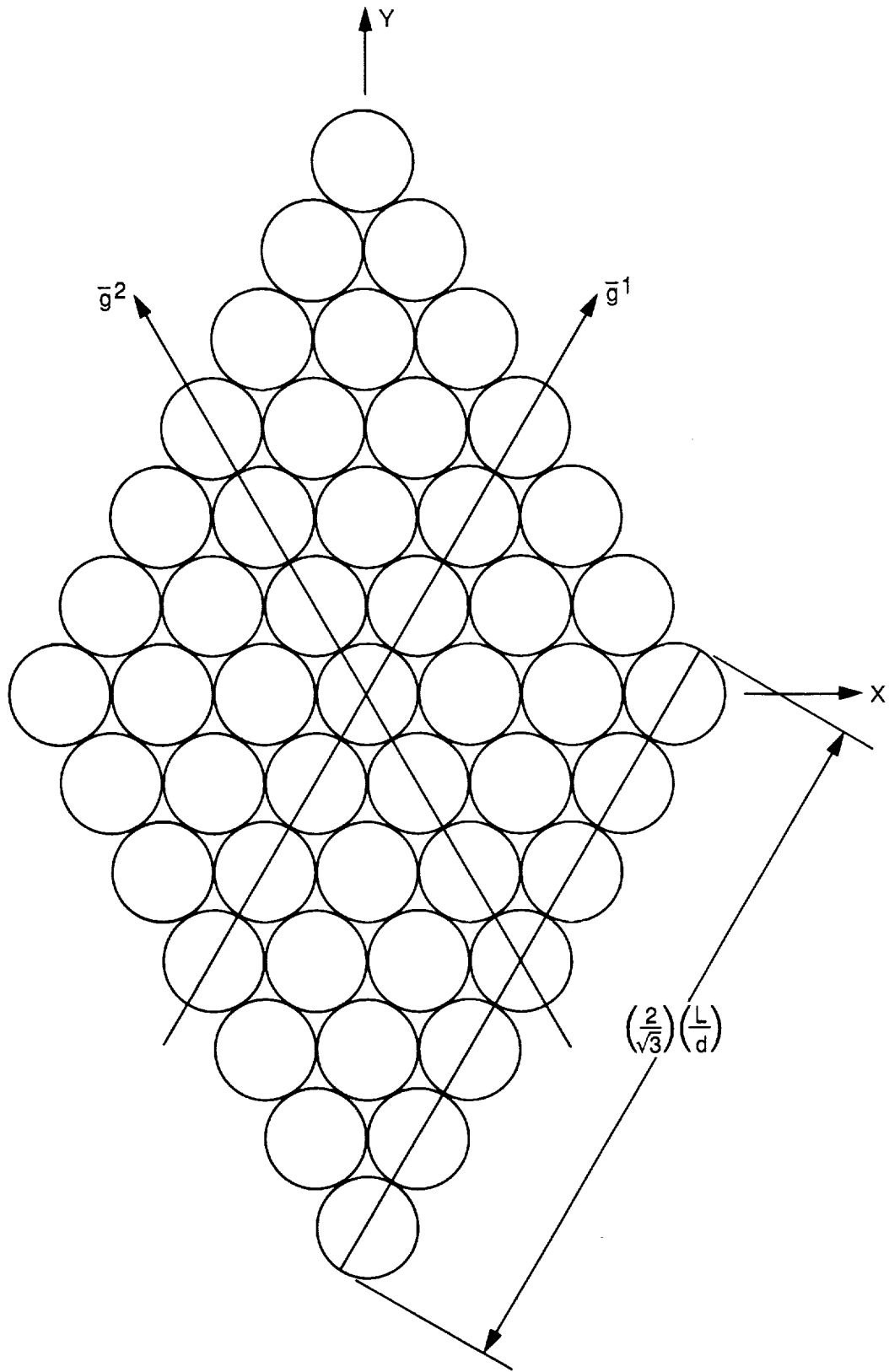$$s \leq 1 \qquad (3.29)$$

Fig. 4. The tiled aperture plane.

Fig. 5. A tentative configuration of aperture-plane data points.
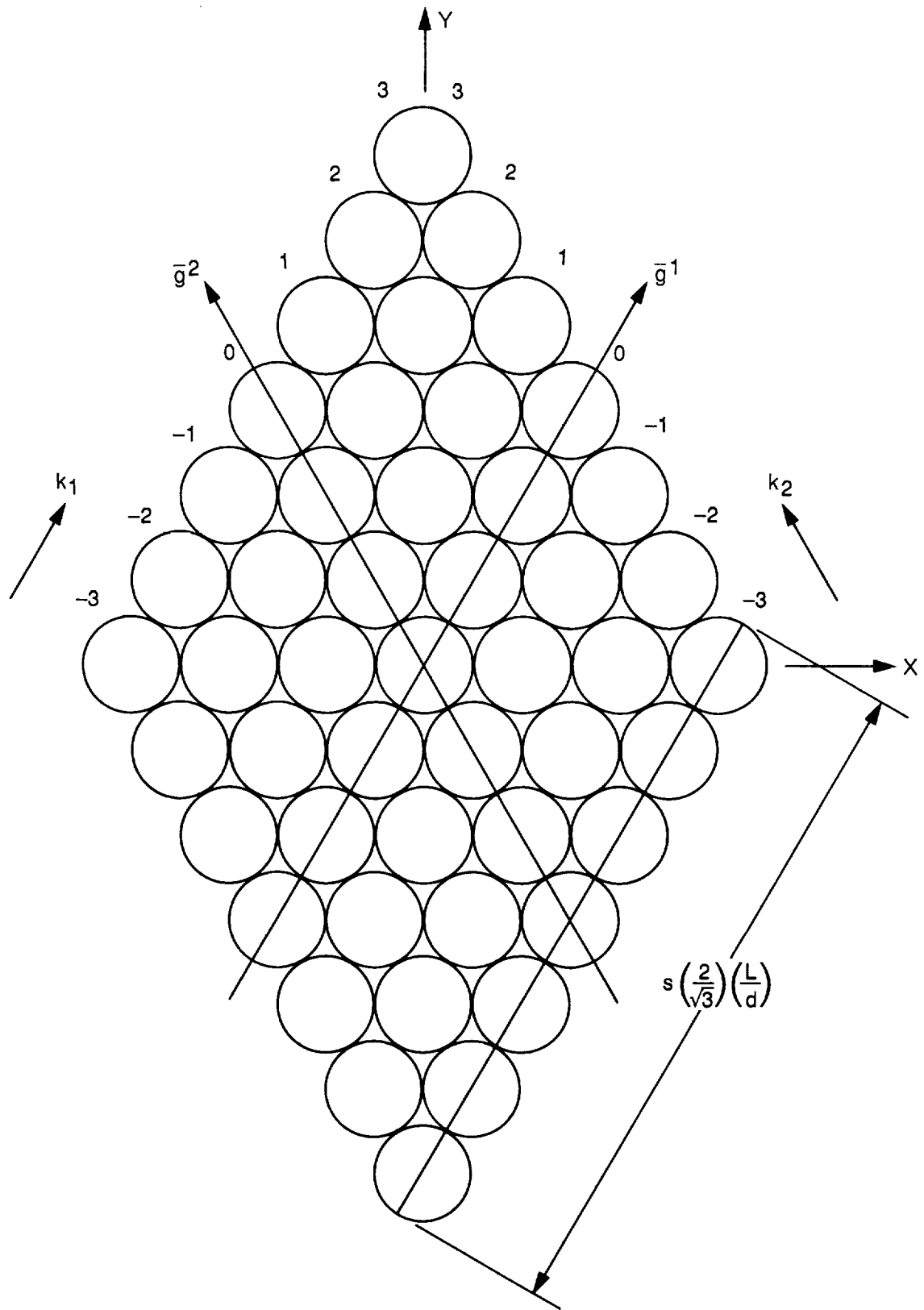
Fig. 6. The adopted configuration of aperture-plane data points.

The importance of the scaling parameter s becomes obvious when we superimpose the circular boundary of the antenna aperture on the tiled aperture plane of Fig. 4. If this circle lies totally within the origin rhombus, then the periodicity of the computed transform becomes a purely academic issue and does not affect us. However, some of the K computed points will fall outside the circle in this case and would thus be wasted. Hence, we are better off shrinking the bounding rhombus, that is, adopting $s < 1$ in (3.28).

If the circle extends beyond the origin rhombus, then the periodicity does manifest itself and, consequently, we have to deal with the effects of aliasing. It turns out that the parameters adopted in the simulations discussed in Section IX fall in this category and Fig. 7 shows the periodic structure across the full aperture for this case.

We obviously have here a severe case of aliasing, that is, we may get significant differences between the "true" value of the transform and its computed value. Consider, for example, point A (of Fig. 7). Due to aliasing, the computed transform at this point is the sum of its true transform plus the true transforms at points B_1, B_2, C_1, C_2, D_1 and D_2. We claim, however that, due to the symmetries associated with these points, the computed value still yields the true phase at point A. To see this, denote by $\hat{\bar{R}}_A$ the (normalized) position vector of point A. Apart from an irrelevant real multiplier, the true voltage at point A is given by

$$\hat{V}_A = e^{i(L\bar{k}\cdot\hat{\bar{R}}_A - \omega t)} \tag{3.30}$$

Now, if we contaminate this value by adding to it the true contributions from points B_1 and B_2 (with position vectors $\hat{\bar{R}}_A \pm \bar{p}^1$), we get

$$\hat{V}_{AB} = e^{i(L\bar{k}\cdot\hat{\bar{R}}_A - \omega t)} [1 + 2 \cos(L\bar{k}\cdot\bar{p}^1)] \tag{3.31}$$

and we see that the phase has not been affected.

The same mechanism operates with the other two pairs of points yielding the more general result

$$\hat{V}_{ABCD} = e^{i(L\bar{k}\cdot\hat{\bar{R}}_A - \omega t)} \{1 + 2 \cos(L\bar{k}\cdot\bar{p}^1) + 2 \cos(L\bar{k}\cdot\bar{p}^2)$$
$$+ 2 \cos[L\bar{k}\cdot(\bar{p}^1 - \bar{p}^2)]\} \tag{3.32}$$

and our claim for point A is thus validated.

However, when we consider point $\hat{A}$ (see Fig. 8), we find that the symmetry argument is applicable only to points $\hat{C}_1$ and $\hat{C}_2$. All the other marked points are unpaired and will

20
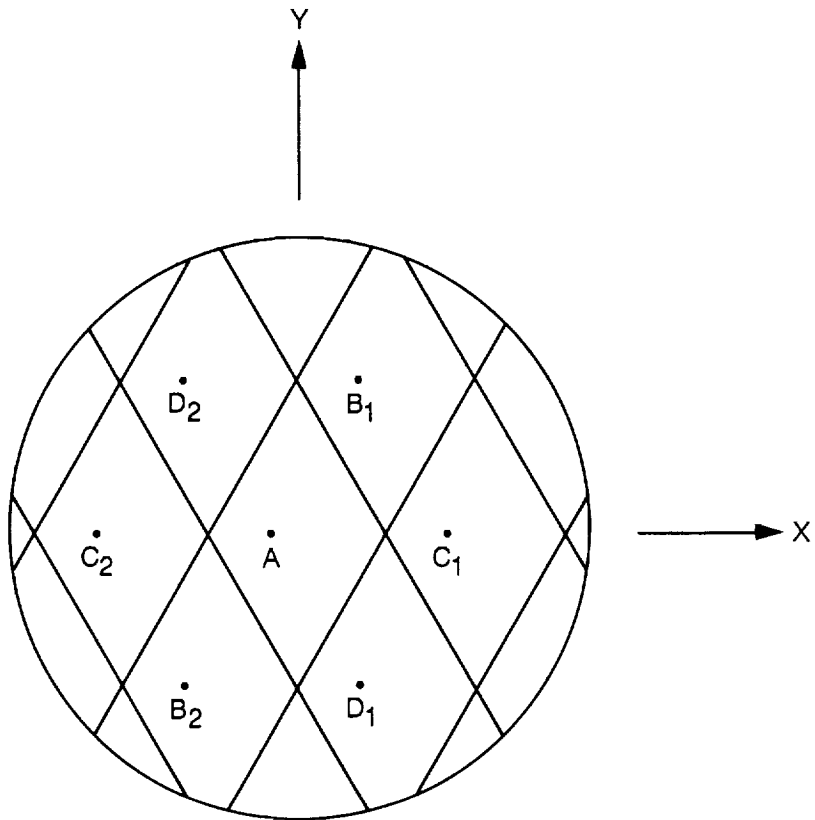
Fig. 7. Example of a point free from aliasing.
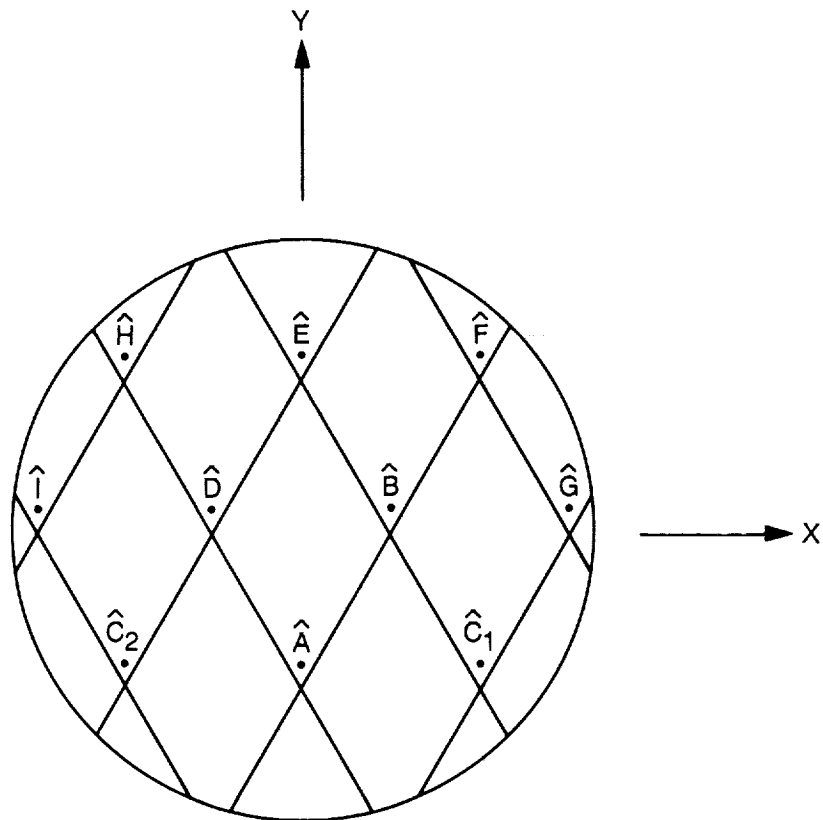
Fig. 8. Example of a point contaminated by aliasing.

thus modify the computed phase at point $\hat{A}$. More generally, we find that all the shaded tiles on the boundary of the aperture in Fig. 9 represent unpaired regions of the aperture and thus cause errors in the corresponding areas of the origin rhombus. When we plot these "forbidden" regions of the origin rhombus, we find that its only region free of aliasing phase errors is the small dotted central area bounded by circular arcs. Trivial computations based on Fig. 9 show that in this particular case we should set $s < 0.156$ to avoid aliasing errors. Note that not all of the unpaired regions have been shaded in Fig. 9. This does not affect our result because the corresponding "forbidden" regions are already accounted for by the regions which have been shaded.

Finally, denoting by $\hat{d}$ the actual (unnormalized) separation of adjacent points, we have (see Fig. 6):

$$\frac{\hat{d}}{L} = \left(\frac{s}{2N+1}\right) \left(\frac{2}{\sqrt{3}}\right) \left(\frac{L}{d}\right) \tag{3.33}$$

This yields the following set of aperture-plane points:

$$\hat{\bar{R}}_k = \hat{\bar{R}}_{k_1,k_2} = (\hat{d}/L) \left[k_1(\bar{g}^1/g^1) + k_2(\bar{g}^2/g^2)\right] \tag{3.34}$$

where k_1, k_2 are integers ranging over the interval $(-N, N)$. Unlike the focal-plane situation (3.18), all such k_i values correspond to actually computed aperture-plane points.

Fig. 9.   Example of an aliasing-free region.

IV. LEAST-SQUARES DETERMINATION OF A WAVEFRONT

In Section II we found that $\bar{n}$, the unit vector pointing to the source, is expressible in terms of the gradient of the phase distribution in the aperture plane (see (2.15)). In the absence of thermal noise and other types of error, three points of the distribution would suffice to determine $\bar{n}$. To 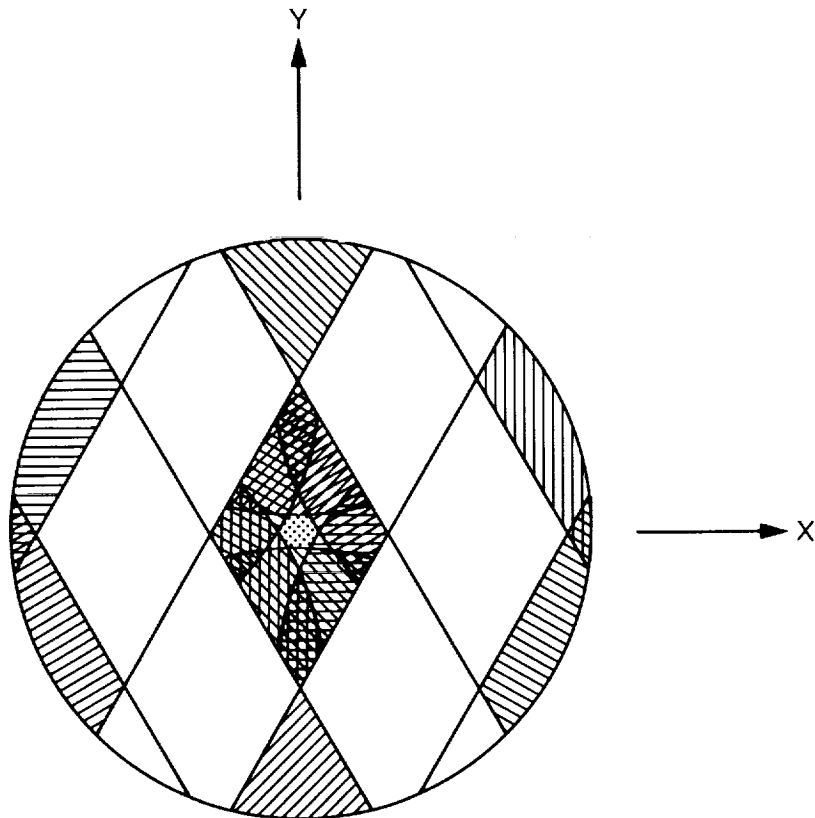partially overcome the effect of noise, we provide more than this minimum number of points and are thus faced with a least-squares determination of $\bar{n}$. However, instead of analyzing a least-squares estimation of a gradient, we prefer to deal with an equivalent—but geometrically more meaningful—problem, namely, the least-squares estimation of the orientation (and location) of a plane from a set of distances to it.

The connection between these two approaches is provided by the following rephrasing of (2.7):

$$\bar{n} = -\bar{\nabla}w \qquad (4.1)$$

where

$$w = \frac{\alpha}{2\pi}\lambda \qquad (4.2)$$

We recognize w as a distance measured in the field of a plane wave along the direction of its propagation. Actually, this is the distance from the point with the measured phase α to the (first) wavefront corresponding to $\alpha = 0$.

The situation is described in Fig. 10 where we show a wavefront (orthogonal to the plane of the figure) at a distance a from an observer located at the origin (point O) and moving towards him. We also show one of a set of K points (P_k) with position vector $\hat{\bar{r}}_k$ and distance w_k from the wavefront. Obviously,

$$w_k = a - \hat{\bar{r}}_k \cdot \bar{n} \qquad (1 \le k \le K) \qquad (4.3)$$

As we shall see shortly, this time it proves convenient to normalize all lengths to $\hat{d}$—the aperture-plane separation of the computed points. Hence, we replace (4.3) by

$$w_k/\hat{d} = (a/\hat{d}) - (\hat{\bar{r}}_k/\hat{d}) \cdot \bar{n} \qquad (1 \le k \le K) \qquad (4.4)$$
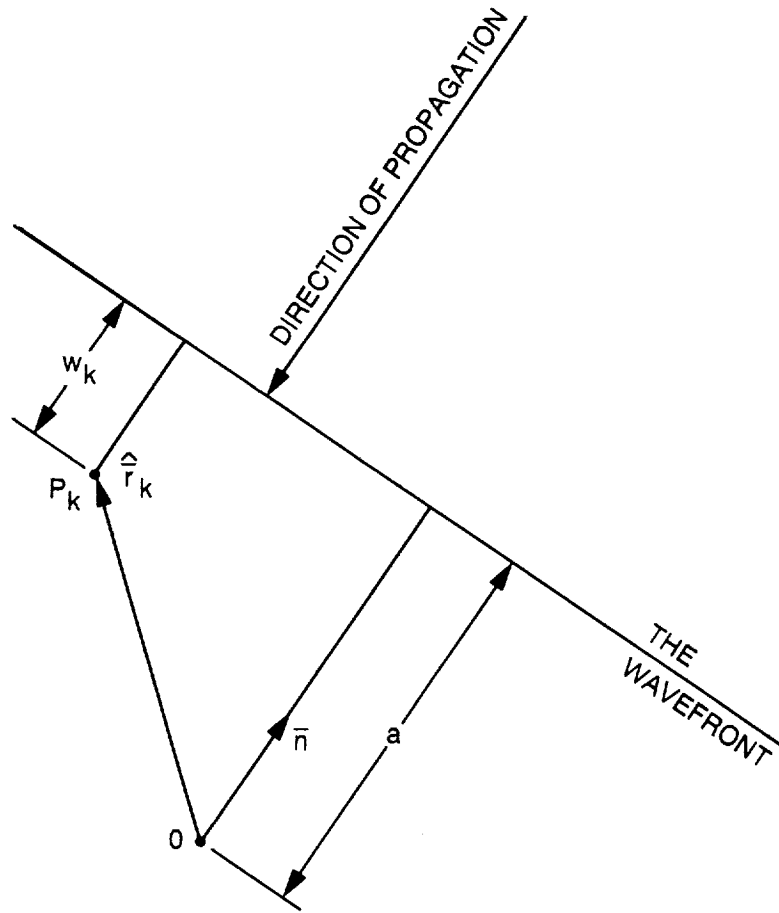
Fig. 10. Wavefront geometry.

Now we infer from (3.34) that

$$\hat{\bar{r}}_k/\hat{d} = \hat{\bar{r}}_{k_1,k_2}/\hat{d} = (\sqrt{3}/2)\left(k_1\bar{g}^1 + k_2\bar{g}^2\right) \qquad (4.5)$$

Substituting (2.13) and (4.5) in (4.4) and applying the $\bar{g}^i$ expansions shown in Fig. 3, we obtain

$$\left[\left(\tfrac{1}{2}\right)(k_2 - k_1)\right]n_1 - \left[(\sqrt{3}/2)(k_2 + k_1)\right]n_2 + (a/\hat{d}) = (w_k/\hat{d}) \qquad (1 \le k \le K) \qquad (4.6)$$

We have here a set of K equations to solve for the three unknowns n_1, n_2 and $(a/\hat{d})$. If the "measured" data (w_k) are totally error-free, then (at least in principle) any three of the above K equations could be used to solve for the three unknowns. If, however, w_k is replaced by its noise-contaminated version $(\hat{w}_k)$, then the absolute equality (=) in (4.6) is replaced by ($\cong$), that is, "equal in the least-squares sense" [3], and (4.6) is transformed into

$$\left[\left(\tfrac{1}{2}\right)(k_2 - k_1)\right]\hat{n}_1 - \left[(\sqrt{3}/2)(k_2 + k_1)\right]\hat{n}_2 + (\hat{a}/\hat{d}) \cong (\hat{w}_k/\hat{d}) \qquad (1 \le k \le K) \qquad (4.7)$$

where $\hat{n}_1$, $\hat{n}_2$ and $\hat{a}$ are the least-squares estimates of n_1, n_2 and a, respectively.

Note that as the number of focal-plane rings (N) increases from 1 to 3, K ranges over the values 9, 25, and 49 (see (3.27)). Thus, we can expect significant noise cancellation in applying least-squares methods to the solution of (4.7). The specific method we choose is the one based on performing a Singular-Value Decomposition (SVD) of the linear operator involved in (4.7). We will show that this approach yields a very simple closed-form solution requiring neither matrix inversion nor any special SVD subroutines. This follows from the very special structure of the linear operator we are dealing with here. In the next section, we investigate this structure and apply it to the solution.

V.   THE STRUCTURE OF THE LINEAR OPERATOR INVOLVED IN THE WAVEFRONT DETERMINATION

We start with a brief overview of the SVD approach to the least-squares problem. Our formulation is quite general with the exception of the dimensionalities and rank, which are those specific to our special case. For a more detailed presentation of this approach see, for example, sections VII and IX of [4].

Assume that we are given the matrix equation

$$A q = p \qquad (5.1)$$

where A is a rectangular $K \times 3$ matrix, and q, p are column matrices of three and K rows, respectively. Given A and a noise-contaminated version of p (denoted $\hat{p}$) , we want to find the (least-squares) best estimate of q (denoted $\hat{q}$) .

The dyadic equivalent of (5.1) is

$$\bar{\bar{A}} \cdot \bar{q} = \bar{p} \qquad (5.2)$$

where $\bar{q}$ is a vector in an abstract three-dimensional space, $\bar{p}$ is a vector in a quite different K-dimensional abstract space, and $\bar{\bar{A}}$ is the dyadic (linear operator) which operates on a vector in the three-dimensional input space to yield a vector in the K-dimensional output space.

The association between a dyadic equation such as (5.2) and the corresponding matrix equation (5.1) always depends on the adoption of two specific orthonormal frames: one $\left(\{\bar{F}_j\}_{j=1}^3 \right)$ which spans the input (three-dimensional) space and one $\left(\{\bar{e}_i\}_{i=1}^K \right)$ which spans the output (K-dimensional) space. More specifically, the column matrix q (with elements q_j) consists of the components of the vector $\bar{q}$ in the $\bar{F}_j$ frame,

$$q_j = \bar{q} \cdot \bar{F}_j \qquad (5.3)$$

the column matrix p (with elements p_i) consists of the components of the vector $\bar{p}$ in the $\bar{e}_i$ frame,

$$p_i = \bar{p} \cdot \bar{e}_i \qquad (5.4)$$

and the rectangular matrix A is the representation of $\bar{\bar{A}}$ in the above two frames, that is (see Appendix A),

$$A_{ij} = \bar{e}_i \cdot \bar{\bar{A}} \cdot \bar{f}_j \tag{5.5}$$

Similarly, $\hat{\bar{q}}$ is the vector whose representation in the $\left(\{\bar{f}_j\}_{j=1}^3\right)$ frame is the column matrix $\hat{q}$ while $\hat{\bar{p}}$ is the vector whose representation in the $\left(\{\bar{e}_i\}_{i=1}^K\right)$ frame is the column matrix $\hat{p}$. The fundamental SVD theorem tells us that our dyadic $\bar{\bar{A}}$ is representable as follows [4]:

$$\bar{\bar{A}} = \sum_{i=1}^{3} s_i \bar{u}_i \bar{v}_i \tag{5.6}$$

where the $\bar{u}_i$'s are K-dimensional orthonormal vectors, the $\bar{v}_i$'s are three-dimensional orthonormal vectors and the s_i's are positive scalars—the singular values of $\bar{\bar{A}}$ (ours is a full-rank case). Dotting both sides of (5.6) with $\bar{v}_j$ (on the right), we see that it is equivalent to

$$\bar{\bar{A}} \cdot \bar{v}_j = s_j \bar{u}_j \tag{5.7}$$

Note that one can always select an arbitrary orthonormal set in the input space, operate on it with $\bar{\bar{A}}$ and get the corresponding output set. In general, however, this output set will not be orthogonal. The special property of the $\bar{v}_i$ set (which makes its determination nontrivial) is that, in addition to being orthogonal itself, the output set it generates in the output space is also orthogonal as implied by (5.7). Incidentally, this provides a simple test for an orthonormal set claimed to be the $\bar{v}_i$ set.

It is a well-known fact that, once the SVD of a linear operator (such as (5.6)) is known, its generalized inverse is immediately available and the solution of the associated least-squares problem becomes trivial (see, for example, sections VII and IX of [4]). In our case, the solution takes the following form:

$$\hat{\bar{q}} = \sum_{i=1}^{3} s_i^{-1} \bar{v}_i \bar{u}_i \cdot \hat{\bar{p}} \tag{5.8}$$

Thus, the main computational effort in finding $\hat{\bar{q}}$ is the decomposition (5.6) rather than the dot products involved in (5.8).

Now we claim that, for the linear operator of (4.7),

$$\bar{v}_i = \bar{f}_i \qquad (1 \leq i \leq 3) \tag{5.9}$$

so that no decomposition is required. In order to verify the validity of this claim, it is convenient to introduce the vectors

$$\bar{z}_i \equiv \bar{\bar{A}} \cdot \bar{f}_i \qquad (1 \leq i \leq 3) \tag{5.10}$$

In view of the preceding discussion, our claim (5.9) is valid if the $\bar{z}_i$ vectors ($1 \leq i \leq 3$) are mutually orthogonal. Note in this context that (see (A-28) and (A-29))

$$\bar{e}_k \cdot \bar{z}_i = \bar{e}_k \cdot \bar{\bar{A}} \cdot \bar{f}_i = A_{ki} \tag{5.11}$$

so that column i of A is the representation of $\bar{z}_i$ in the $\bar{e}_k$ frame. Therefore, an equivalent formulation of the validity condition is that the columns of A be mutually orthogonal.

To facilitate the verification of this condition, let us express the variables of (4.7) in terms of the new variables introduced in this section.

$$\hat{n}_j = \hat{\bar{q}} \cdot \bar{f}_j = \hat{q}_j \qquad (j = 1, 2) \tag{5.12}$$

$$\hat{a}/\hat{d} = \hat{\bar{q}} \cdot \bar{f}_3 = \hat{q}_3 \tag{5.13}$$

$$\hat{w}_k/\hat{d} = \hat{\bar{p}} \cdot \bar{e}_k \tag{5.14}$$

$$\bar{e}_k \cdot \bar{z}_1 = A_{k,1} = (1/2)(k_2 - k_1) \tag{5.15}$$

30

$$\bar{e}_k \cdot \bar{z}_2 = A_{k,2} = -(\sqrt{3}/2)(k_2 + k_1) \tag{5.16}$$

$$\bar{e}_k \cdot \bar{z}_3 = A_{k,3} = 1 \tag{5.17}$$

We proceed now to check the three dot products in the $\{\bar{z}_i\}_{i=1}^3$ set.

$$\bar{z}_1 \cdot \bar{z}_3 = (1/2) \sum_{k_1, k_2 = -N}^{N} (k_2 - k_1) = 0 \tag{5.18}$$

$$\bar{z}_2 \cdot \bar{z}_3 = -(\sqrt{3}/2) \sum_{k_1, k_2 = -N}^{N} (k_2 + k_1) = 0 \tag{5.19}$$

$$\bar{z}_1 \cdot \bar{z}_2 = -(\sqrt{3}/4) \sum_{k_1, k_2 = -N}^{N} (k_2^2 - k_1^2) = 0 \tag{5.20}$$

We conclude that (5.9) is valid and proceed to determine the remaining SVD parameters required for the solution (5.8). First, we infer from (5.7) that a corollary of (5.9) is

$$\bar{z}_i = s_i \bar{u}_i \tag{5.21}$$

and, consequently,

$$s_i = z_i \tag{5.22}$$

Thus, we can rewrite the solution (5.8) as follows:

$$\hat{\bar{q}} = \sum_{i=1}^{3} z_i^{-2} \bar{\bar{f}}_i \bar{z}_i \cdot \hat{\bar{p}} \tag{5.23}$$

But, in view of (5.12) and (5.13), this means

$$\hat{n}_i = \frac{\bar{z}_i \cdot \hat{p}}{z_i^2} \qquad (i = 1, 2) \qquad (5.24)$$

$$\hat{a}/\hat{d} = \frac{\bar{z}_3 \cdot \hat{p}}{z_3^2} \qquad (5.25)$$

Of the three squared magnitudes appearing here, z_3^2 is obtained trivially as

$$s_3^2 = z_3^2 = K \qquad (5.26)$$

In computing the other two, we use the following well-known summation formula (Series No. (19) of [5])

$$\sum_{m=1}^{M} m^2 = (1/6) M(M + 1)(2M + 1) \qquad (5.27)$$

which yields (using (3.15))

$$
\begin{aligned}
s_1^2 = z_1^2 &= (1/4) \sum_{k_1, k_2 = -N}^{N} (k_1^2 + k_2^2 - 2k_1 k_2) \\
&= (1/6)(2N + 1)^2 N(N + 1) = (1/18) K(J - 1) \qquad (5.28)
\end{aligned}
$$

$$
\begin{aligned}
s_2^2 = z_2^2 &= (3/4) \sum_{k_1, k_2 = -N}^{N} (k_1^2 + k_2^2 + 2k_1 k_2) \\
&= (1/2)(2N + 1)^2 N(N + 1) = (1/6) K(J - 1) \qquad (5.29)
\end{aligned}
$$

It should be pointed out that we have deviated here from the standard SVD convention regarding the indexing of the singular values, namely, $s_i \geq s_{i+1}$. Our hand is forced here because the ranking of the s_i's (by value) varies with N.

Applying the values obtained here, we get the very simple final results:

$$\hat{n}_1 = \{9 / [K(J - 1)\hat{d}]\} \sum_{k_1, k_2 = -N}^{N} (k_2 - k_1) \hat{w}_{k_1, k_2} \qquad (5.30)$$

$$\hat{n}_2 = -\{3\sqrt{3} / [K(J - 1)\hat{d}]\} \sum_{k_1, k_2 = -N}^{N} (k_2 + k_1) \hat{w}_{k_1, k_2} \qquad (5.31)$$

$$\hat{a} = (1/K) \sum_{k=1}^{K} \hat{w}_k \qquad (5.32)$$

Note that the equations for $\hat{n}_1$, $\hat{n}_2$ may be expressed in terms of templates applied to the $\hat{w}_{k_1, k_2}$ distances, that is, each $\hat{w}_{k_1, k_2}$ is multiplied by the template value corresponding to its position in the aperture plane and the estimated parameter is the sum of these template-weighted distances. These templates are shown in Figs. 11 and 12, which display very distinct simple patterns.

These patterns suggest that, as long as we restrict the $\hat{\bar{R}}_k$ points to a small fraction of the basic-period cell, placing the K aperture-plane points inside a square (rather than a rhombus) is probably an equally good alternative to the rhombic arrangement.

Though we now have a simple algorithm which yields the best (least-squares) estimate of the DOA, its practical significance is rather minimal unless we complement it with a second algorithm which will provide us with information on the uncertainties associated with this estimate. It turns out that this is a much more challenging undertaking; it will occupy us in the next three sections.

In preparation for this task, we point out two results which, though of no direct interest, are required in the development of the second algorithm. The first one is the (trivial) value for the estimate of a —the distance to the reference wavefront (5.32). The second one is the complete vectorial representation of the least-squares estimate of $\bar{n}$, which we denote $\hat{\bar{n}}$. Obviously,

$$\hat{\bar{n}} = \hat{n}_1 \bar{h}_1 + \hat{n}_2 \bar{h}_2 + \left(\sqrt{1 - \hat{n}_1^2 - \hat{n}_2^2} \right) \bar{h}_3 \qquad (5.33)$$

where $\{\bar{h}_j\}_{j=1}^3$ is the antenna-fixed frame (see Fig. 1).

$$\hat{n}_1 = \frac{9}{K(J-1)} \sum_{k=1}^{K} \left(\frac{\hat{w}_k}{\hat{d}} \right) \text{ (TEMPLATE VALUE FOR POINT k)}$$
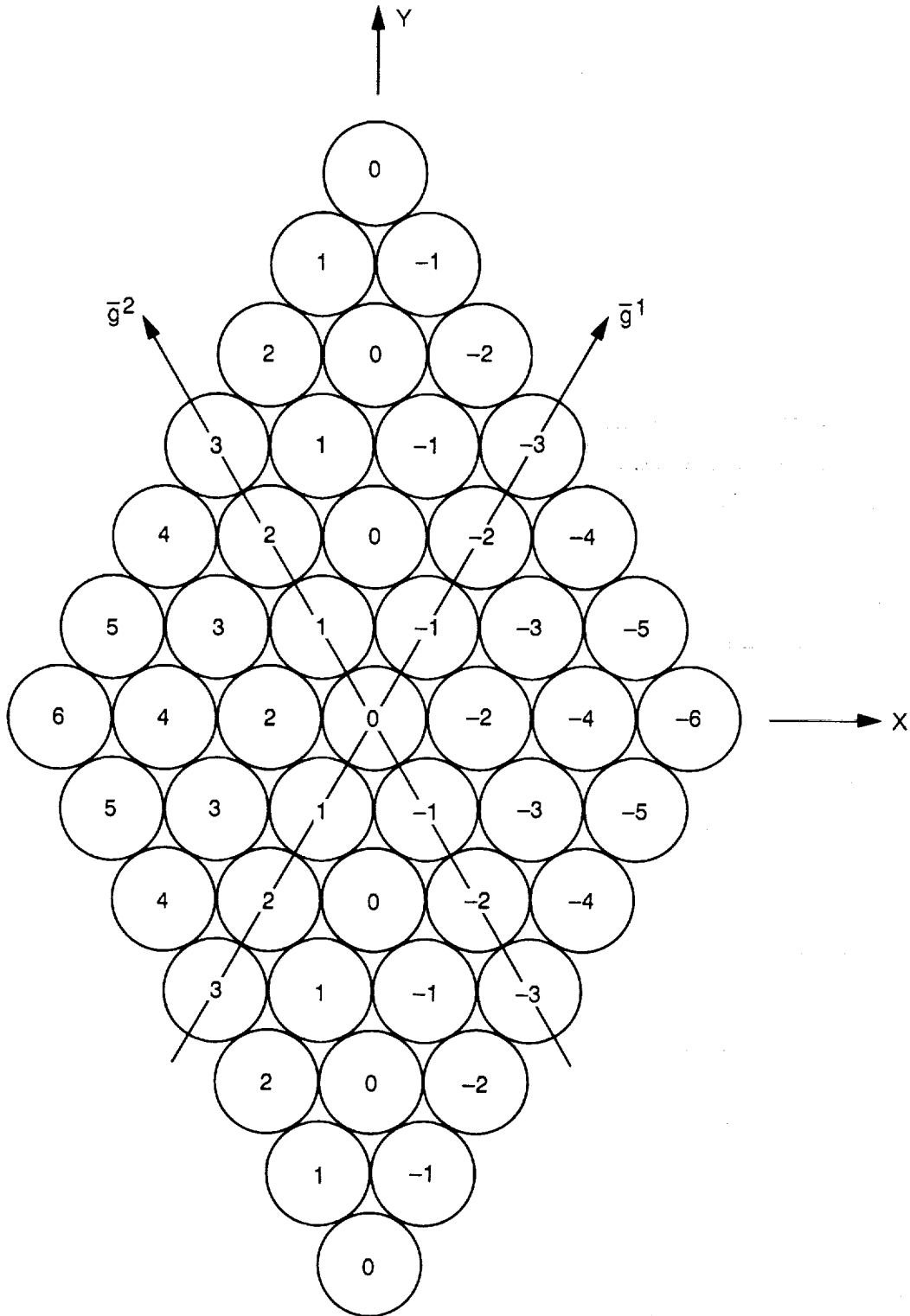

Fig. 11. The template for estimating n_1.

$$\hat{n}_2 = \frac{3\sqrt{3}}{K(J-1)} \sum_{k=1}^{K} \left(\frac{\hat{w}_k}{\hat{d}}\right) \text{(TEMPLATE VALUE FOR POINT k)}$$
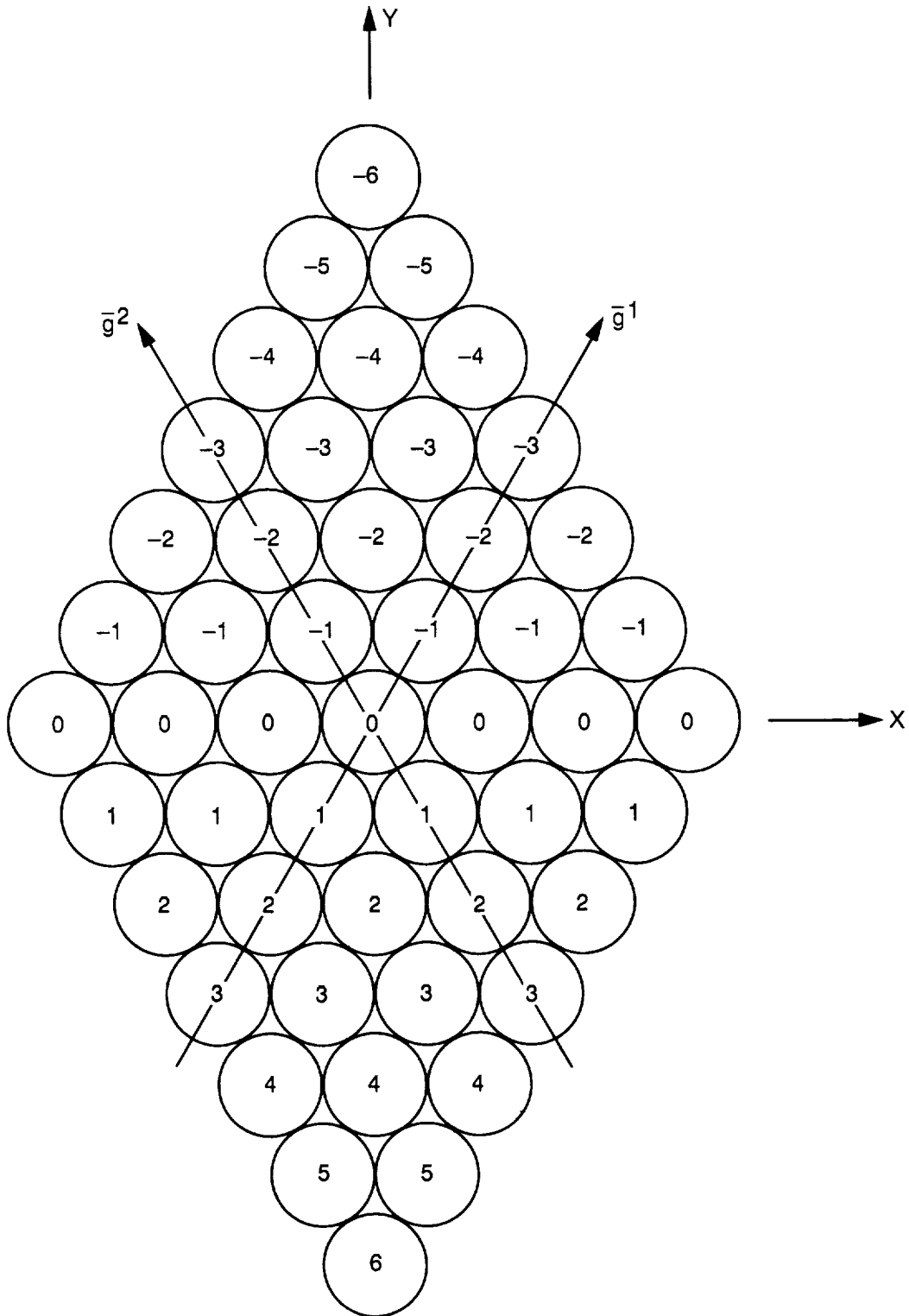
Fig. 12. The template for estimating n_2.

VI. CONFIDENCE BOUNDS FOR THE DOA DETERMINATION (PRELIMINARY CONSIDERATIONS)

The DOA may be represented as a point (D) on the unit sphere (the intersection of $\bar{n}$ with the unit sphere). Having solved for $\hat{\bar{n}}$, we are now in possession of the location of point $\hat{D}$, which is the best (least-squares) estimate of point D. Now we want to know the reliability of this estimate. Specifically, we want to know the size (and shape) of the region around $\hat{D}$ where the true point D is likely to be found with a prescribed probability (say, 99%). To attain that goal, we have to start with the statistical properties of the least-squares vector from which $\hat{\bar{n}}$ is derived, namely, the least-squares solution vector $\hat{\bar{q}}$.

First, we recall that the very nature of the least-squares processing implies

$$E\{\hat{\bar{q}}\} = \bar{q} \qquad (6.1)$$

where E is the expectation operator. Expressing this in terms of the first-order central moment of $\hat{\bar{q}}$, we have

$$E\{\hat{\bar{q}} - \bar{q}\} = \bar{0} \qquad (6.2)$$

To attain our goal, we have to evaluate the second-order central moment of $\hat{\bar{q}}$, namely,

$$\bar{\bar{Y}} = E\{(\hat{\bar{q}} - \bar{q})(\hat{\bar{q}} - \bar{q})\} \qquad (6.3)$$

Note that $\bar{\bar{Y}}$ is the covariance dyadic whose relationship to the (more familiar) covariance matrix Y (with elements Y_{ij}) is expressible as follows:

$$\bar{\bar{Y}} = \sum_{i,j=1}^{3} Y_{ij} \bar{f}_i \bar{f}_j \qquad (6.4)$$

To evaluate $\bar{\bar{Y}}$, we have to complement the expression for $\hat{\bar{q}}$ (5.8) with a similar expression for $\bar{q}$, namely,

$$\bar{q} = \sum_{i=1}^{3} s_i^{-1} \bar{v}_i \bar{u}_i \cdot \bar{p} \qquad (6.5)$$

This is simply the special case of (5.8) in which there is no noise corrupting the data vector $\bar{p}$. The only validity condition (for both equations) is the requirement that $\bar{\bar{A}}$ should have full rank (3) or, equivalently, that all three singular values be nonzero. This is, indeed, the case as evidenced by (5.26), (5.28), and (5.29).

Subtracting (6.5) from (5.8) and recalling (5.9), we obtain

$$\hat{\bar{q}} - \bar{q} = \sum_{i=1}^{3} s_i^{-1} \bar{f}_i \bar{u}_i \cdot (\hat{\bar{p}} - \bar{p}) \tag{6.6}$$

The vector $(\hat{\bar{p}} - \bar{p})$ appearing here represents the aperture-plane noise, which is the cause of the uncertainty in $\bar{q}$. Let us denote

$$\hat{\bar{p}} - \bar{p} \equiv \bar{\zeta} \tag{6.7}$$

In terms of components in the $\bar{e}_i$ frame, this means (see (5.14))

$$\zeta_i \equiv \bar{e}_i \cdot \bar{\zeta} = (\hat{w}_i - w_i)/\hat{d} \tag{6.8}$$

Now we apply (6.6) and (6.7) to the evaluation of $\bar{\bar{Y}}$ (6.3), getting

$$\bar{\bar{Y}} = \sum_{i,j=1}^{3} s_i^{-1} s_j^{-1} \bar{f}_i \bar{u}_i \cdot E\{\bar{\zeta}\bar{\zeta}\} \cdot \bar{u}_j \bar{f}_j \tag{6.9}$$

In view of (6.4) and (5.21), this means

$$Y_{ij} = \frac{\bar{z}_i \cdot E\{\bar{\zeta}\bar{\zeta}\} \cdot \bar{z}_j}{z_i^2 \, z_j^2} \tag{6.10}$$

The only unknown entity here is $E\{\bar{\zeta}\bar{\zeta}\}$, which is the covariance dyadic of the noise contaminating the data vector of the least-squares problem. In the usual least-squares application, one assumes the corresponding covariance matrix to be diagonal and, based on that, simplifies the Y_{ij} expression. In our case, the situation is quite different because the noise is injected into the system not in the

37

aperture plane where the least-squares algorithm is applied, but in the focal plane. Thus, while the focal-plane noise is, indeed, uncorrelated, the processing that generates the aperture-plane distribution leads to highly correlated aperture-plane noise. Our immediate goal, then, is the evaluation of the covariance dyadic of this noise, which we take up in the next section.

VII. THE COVARIANCE OF THE APERTURE-PLANE NOISE

As outlined in the last section, our task here is to evaluate the dyadic $E\{\bar{\zeta}\,\bar{\zeta}\}$ or, equivalently, the set of scalars $E\{\zeta_k\zeta_m\}$ for all k, m. The parameters which determine ζ_k are illustrated in Fig. 13, which shows the k-th noise-free complex voltage with magnitude a_k and phase angle α_k. This voltage is contaminated by the added complex noise voltage $\hat{N}_k$ (with magnitude N_k), which leads to the phase error η_k as shown. From (6.8) and (4.2), we infer

$$\zeta_k = (\eta_k/2\pi)\,(\lambda/\hat{d}) \tag{7.1}$$

so that

$$E\{\zeta_k\zeta_m\} = [\lambda/(2\pi\hat{d})]^2\,E\{\eta_k\eta_m\} \tag{7.2}$$

and we concentrate our attention initially on $E\{\eta_k\eta_m\}$.

Though we are dealing here with complex scalars, it is convenient to regard them as three-dimensional vectors (over the field of reals) in a space with an orthonormal basis $\{\bar{c}_i\}_{i=1}^3$, as indicated in Fig. 13. Note that, in this isomorphism, the vector $\bar{c}_1$ is isomorphic to the real number 1, the vector $\bar{c}_2$ is isomorphic to the imaginary number i, and the vector $\bar{N}_k$ is isomorphic to the complex number $\hat{N}_k$. The motivation for establishing this isomorphism with three-dimensional vectors rather than two-dimensional vectors will become clear shortly.

The computation of $E\{\eta_k\eta_m\}$ in the general case presents serious mathematical difficulties. However, if we constrain our analysis to the practically significant case of large SNR, the problem becomes manageable. We propose, then, to pursue this initial analysis under the constraint

$$N_k \ll a_k \tag{7.3}$$

This means that, even in the worst-case orientation, we have

$$\eta_k \approx \tan\eta_k \tag{7.4}$$

so that

$$\eta_k \approx \frac{N_k\sin\psi_k}{a_k + N_k\cos\psi_k} \approx \frac{N_k\sin\psi_k}{a_k} \tag{7.5}$$
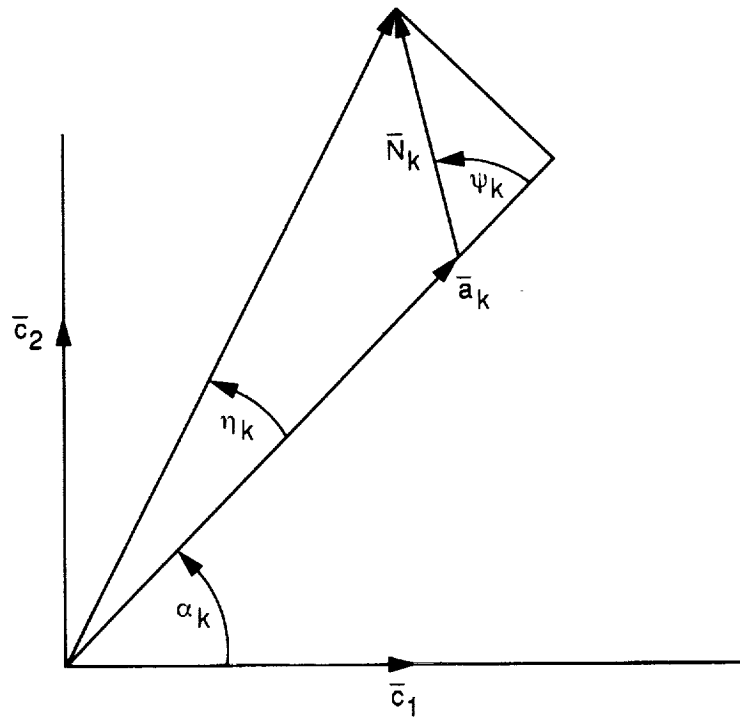
Fig. 13. The aperture-plane phase-error geometry.

But this is expressible as a vector product, namely,

$$\eta_k \bar{c}_3 = \frac{\bar{a}_k \times \bar{N}_k}{a_k^2} \tag{7.6}$$

Hence, we can express the product $(\eta_k \eta_m)$ as a dot product:

$$\eta_k \eta_m = \left(\frac{\bar{a}_k \times \bar{N}_k}{a_k^2} \right) \cdot \left(\frac{\bar{a}_m \times \bar{N}_m}{a_m^2} \right) \tag{7.7}$$

Applying now elementary vector manipulations, we reduce this to

$$(a_k a_m)^2 \, (\eta_k \eta_m) = (\bar{a}_k \cdot \bar{a}_m)(\bar{N}_k \cdot \bar{N}_m) - \bar{a}_m \cdot \bar{N}_k \bar{N}_m \cdot \bar{a}_k \tag{7.8}$$

and finally,

$$(a_k a_m)^2 \, E\{\eta_k \eta_m\} = (\bar{a}_k \cdot \bar{a}_m) E\{\bar{N}_k \cdot \bar{N}_m\} - \bar{a}_m \cdot E\{\bar{N}_k \bar{N}_m\} \cdot \bar{a}_k \tag{7.9}$$

To evaluate the expectations appearing in (7.9), we have to trace the evolution of $\bar{N}_k$ from the noise generated in the focal plane. The motivation here is that we can safely make reasonable assumptions concerning the statistical properties of the noise added to the focal-plane horns' outputs. Let the complex noise added to the output of the $j-th$ horn be $\hat{n}_j$. Applying (3.17), we find that

$$\hat{N}_k = \sum_{j=1}^{J} e^{i\pi \bar{R}_j \cdot (2\hat{R}_k - \bar{R}_j)} \hat{n}_j \tag{7.10}$$

Applying now the isomorphism introduced earlier, we may represent the complex noise $\hat{n}_j$ by the real vector $\bar{n}_j$ and (7.10) will have to be expressed in terms of rotation operators. Specifically, the complex scalar $e^{i\beta}\hat{n}_j$ is isomorphic to the real vector $\bar{\bar{T}}(\bar{c}_3, \beta) \cdot \bar{n}_j$ where

$$\overline{\overline{T}}(\overline{\rho}, \beta) = \text{a rotation dyadic which, on operating to the right, rotates a}$$

vector by the angle β about an axis given by the unit vector $\overline{\rho}$ (7.11)

Though there is a relatively simple explicit formulation of $\overline{\overline{T}}(\overline{\rho}, \beta)$ in terms of its $\overline{\rho}$, β parameters, we do not need this here. For our derivations, it is sufficient to realize that $\overline{\overline{T}}(\overline{\rho}, \beta)$ is a length-preserving linear operator (represented by an orthogonal matrix) and thus satisfies

$$\overline{\overline{T}}^{-1}(\overline{\rho}, \beta) = \overline{\overline{T}}(\overline{\rho}, -\beta) = \overline{\overline{\overline{T}}}(\overline{\rho}, \beta) \qquad (7.12)$$

Applying these ideas to (7.10), we get

$$\overline{N}_k = \sum_{j=1}^{J} \overline{\overline{T}}[\overline{c}_3, \pi \overline{R}_j \cdot (2\hat{\overline{R}}_k - \overline{R}_j)] \cdot \overline{n}_j \qquad (7.13)$$

To simplify the notation, we denote

$$\overline{\overline{T}}_{kj} \equiv \overline{\overline{T}}[\overline{c}_3, \pi \overline{R}_j \cdot (2\hat{\overline{R}}_k - \overline{R}_j)] \qquad (7.14)$$

so that

$$\overline{N}_k = \sum_{j=1}^{J} \overline{\overline{T}}_{kj} \cdot \overline{n}_j = \sum_{j=1}^{J} \overline{n}_j \cdot \overline{\overline{\overline{T}}}_{kj} \qquad (7.15)$$

Now we apply (7.15) to evaluate the two expectations appearing on the right of (7.9). Specifically,

$$E\{\overline{N}_k \cdot \overline{N}_m\} = \sum_{j,s=1}^{J} E\{\overline{n}_j \cdot \overline{\overline{\overline{T}}}_{kj} \cdot \overline{\overline{T}}_{ms} \cdot \overline{n}_s\} \qquad (7.16)$$

42

$$E\{\bar{N}_k \bar{N}_m\} = \sum_{j,s=1}^{J} \bar{\bar{T}}_{kj} \cdot E\{\bar{n}_j \bar{n}_s\} \cdot \bar{\bar{T}}_{ms} \qquad (7.17)$$

To proceed further, we have to utilize now the statistical properties of $\bar{n}_j$, the noise component of the output of the j-th horn. We adopt the following two assumptions concerning these noise components:

(1)   All noise components (in-phase and quadrature) of all the horns' outputs are totally uncorrelated.

(2)   The in-phase and quadrature noise components of the j-th horn have the same power (σ_j^2).

To facilitate the mathematical formulation of these assumptions, we express $\bar{n}_j$ in terms of its components as follows:

$$\bar{n}_j = \sum_{u=1}^{2} n_{ju} \bar{c}_u \qquad (7.18)$$

Our two assumptions are equivalent to

$$E\{n_{ju} n_{sv}\} = \sigma_j^2 \delta_{js} \delta_{uv} \qquad (7.19)$$

We apply this equation to (7.17) first, starting with the evaluation of

$$E\{\bar{n}_j \bar{n}_s\} = \sum_{u,v=1}^{2} E\{n_{ju} n_{sv}\} \bar{c}_u \bar{c}_v$$

$$= \sigma_j^2 \delta_{js} \sum_{u=1}^{2} \bar{c}_u \bar{c}_u$$

$$= \sigma_j^2 \delta_{js} (\bar{\bar{I}} - \bar{c}_3 \bar{c}_3) \qquad (7.20)$$

where $\bar{\bar{I}}$ is the identity dyadic in the space spanned by the frame $\{\bar{c}_i\}_{i=1}^{3}$. Using this result in (7.17), we obtain

$$E\{\bar{N}_k \bar{N}_m\} = \sum_{j=1}^{J} \sigma_j^2 \bar{\bar{T}}_{kj} \cdot (\bar{\bar{I}} - \bar{c}_3 \bar{c}_3) \cdot \bar{\bar{T}}_{mj}$$

$$= \left(\sum_{j=1}^{J} \sigma_j^2 \bar{\bar{T}}_{kj} \cdot \bar{\bar{T}}_{mj} \right) - \bar{c}_3 \bar{c}_3 \sum_{j=1}^{J} \sigma_j^2 \qquad (7.21)$$

We have used here the obvious fact that a rotation operator applied to the rotation axis $(\bar{c}_3)$ leaves it unchanged. Resorting now to the $\bar{\bar{T}}_{kj}$ definition (7.14), we can replace the product of the two rotation dyadics we have here by a single rotation dyadic as follows:

$$\bar{\bar{T}}_{kj} \cdot \bar{\bar{T}}_{mj} = \bar{\bar{T}}[\bar{c}_3, -\pi \bar{R}_j \cdot (2\hat{\bar{R}}_m - \bar{R}_j) + \pi \bar{R}_j \cdot (2\hat{\bar{R}}_k - \bar{R}_j)]$$

$$= \bar{\bar{T}}[\bar{c}_3, 2\pi(\hat{\bar{R}}_k - \hat{\bar{R}}_m) \cdot \bar{R}_j] \qquad (7.22)$$

The angular argument appearing here turns out to be an important system parameter. We denote it γ_{kmj}, that is,

$$\gamma_{kmj} \equiv 2\pi(\hat{\bar{R}}_k - \hat{\bar{R}}_m) \cdot \bar{R}_j \qquad (7.23)$$

so that

$$\bar{\bar{T}}_{kj} \cdot \bar{\bar{T}}_{mj} = \bar{\bar{T}}(\bar{c}_3, \gamma_{kmj}) \qquad (7.24)$$

and, finally,

$$E\{\bar{N}_k \bar{N}_m\} = \left[\sum_{j=1}^{J} \sigma_j^2 \bar{\bar{T}}(\bar{c}_3, \gamma_{kmj}) \right] - \bar{c}_3 \bar{c}_3 \sum_{j=1}^{J} \sigma_j^2 \qquad (7.25)$$

44

We turn now to the other expectation, namely (7.16). Applying (7.19) to it, we obtain

$$E\{\bar{N}_k \cdot \bar{N}_m\} = \sum_{j,s=1}^{J} \sum_{u,v=1}^{2} E\{n_{ju}n_{sv}\}\bar{c}_u \cdot \overset{\cong}{T}_{kj} \cdot \overset{=}{T}_{ms} \cdot \bar{c}_v$$

$$= \sum_{j=1}^{J} \sum_{u=1}^{2} \sigma_j^2 \bar{c}_u \cdot \overset{\cong}{T}_{kj} \cdot \overset{=}{T}_{mj} \cdot \bar{c}_u \qquad (7.26)$$

Repeating the manipulations involved in (7.22), we get in the present case

$$\overset{\ast}{T}_{kj} \cdot \overset{=}{T}_{mj} = \overset{=}{T}(\bar{c}_3, \ -\gamma_{kmj}) \qquad (7.27)$$

so that

$$E\{\bar{N}_k \cdot \bar{N}_m\} = \sum_{j=1}^{J} \sum_{u=1}^{2} \sigma_j^2 \ [\bar{c}_u \cdot \overset{=}{T}(\bar{c}_3, \ -\gamma_{kmj}) \cdot \bar{c}_u] \qquad (7.28)$$

Having computed the two required expectations, we go back now to (7.9) and substitute (7.25) and (7.28) in it. Recalling that the vectors $\bar{a}_k, \bar{a}_m$ are orthogonal to $\bar{c}_3$, the result is

$$(a_k a_m)^2 E\{\eta_k \eta_m\} = (\bar{a}_k \cdot \bar{a}_m) \sum_{j=1}^{J} \sum_{u=1}^{2} \sigma_j^2 \ [\bar{c}_u \cdot \overset{=}{T}(\bar{c}_3, \ -\gamma_{kmj}) \cdot \bar{c}_u]$$

$$- \sum_{j=1}^{J} \sigma_j^2 \ [\bar{a}_m \cdot \overset{=}{T}(\bar{c}_3, \ \gamma_{kmj}) \cdot \bar{a}_k] \qquad (7.29)$$

We have here two quadratic forms involving a rotation dyadic. Since all the vectors in these quadratic forms are orthogonal to the rotation axis $(\bar{c}_3)$, the geometric interpretation of these quadratic forms is very simple. Consider the second quadratic form while reviewing Fig. 13. Vector $\bar{a}_k$ has magnitude a_k and forms an angle α_k with $\bar{c}_1$. Operating on it with $\overset{=}{T}(\bar{c}_3, \gamma_{kmj})$ yields a vector of magnitude a_k forming an angle $(\alpha_k + \gamma_{kmj})$ with $\bar{c}_1$. Finally, dotting this vector with $\bar{a}_m$ yields the result

$$\bar{a}_m \cdot \bar{\bar{T}}(\bar{C}_3, \gamma_{kmj}) \cdot \bar{a}_k = a_k a_m \cos(\alpha_k - \alpha_m + \gamma_{kmj}) \tag{7.30}$$

Applying the same reasoning to the other quadratic form yields

$$\sum_{u=1}^{2} \bar{c}_u \cdot \bar{\bar{T}}(\bar{C}_3, -\gamma_{kmj}) \cdot \bar{c}_u = 2 \cos \gamma_{kmj} \tag{7.31}$$

These results simplify (7.29), yielding (after some trivial trigonometric manipulations)

$$(a_k a_m) E\{\eta_k \eta_m\} = \cos(\alpha_k - \alpha_m) \sum_{j=1}^{J} \sigma_j^2 \cos \gamma_{kmj} + \sin(\alpha_k - \alpha_m) \sum_{j=1}^{J} \sigma_j^2 \sin \gamma_{kmj} \tag{7.32}$$

The two sums appearing here are system parameters involving geometric and noise entities, but these sums are independent of the signal. For a given system configuration, they are constants which have to be evaluated only once. We introduce now the following symbols to represent these sums:

$$\hat{C}_{km} = \sum_{j=1}^{J} \sigma_j^2 \cos[2\pi(\hat{\bar{R}}_k - \hat{\bar{R}}_m) \cdot \bar{R}_j] \tag{7.33}$$

$$\hat{S}_{km} = \sum_{j=1}^{J} \sigma_j^2 \sin[2\pi(\hat{\bar{R}}_k - \hat{\bar{R}}_m) \cdot \bar{R}_j] \tag{7.34}$$

Note that we have replaced γ_{kmj} by its value (7.23) to bring out the explicit dependence on the geometry. In terms of these symbols, our final result is

$$E\{\eta_k \eta_m\} = [1/(a_k a_m)][\hat{C}_{km} \cos(\alpha_k - \alpha_m) + \hat{S}_{km} \sin(\alpha_k - \alpha_m)] \tag{7.35}$$

This formulation allows us to handle the case where different amplifier types are used for different horns or to accommodate variations in the noise performance of supposedly identical amplifiers.

46

One cost-effective design would use identical (cheaper, higher-noise) amplifiers for all horns except the central one [6]. In this case, we can take advantage of the symmetry of the distribution of the horns in the focal plane. Note that all horns other than the central one occur in pairs at locations $\pm \bar{R}_j$ (see Fig. 2), and thus their contributions to $\hat{S}_{km}$ cancel out. The central horn makes no contribution either since the corresponding $\bar{R}_j$ is zero. Thus, $\hat{S}_{km}$ vanishes in this case and we get the simpler result:

$$E\{\eta_k \eta_m\} = \hat{C}_{km} \frac{\cos(\alpha_k - \alpha_m)}{a_k a_m} \tag{7.36}$$

Finally, if all the amplifiers are identical with the common noise power σ^2, then

$$\hat{C}_{km} = \sigma^2 C_{km} \tag{7.37}$$

where C_{km} is a purely geometric factor given by

$$C_{km} = \sum_{j=1}^{J} \cos[2\pi(\hat{\bar{R}}_k - \hat{\bar{R}}_m) \cdot \bar{R}_j] \tag{7.38}$$

and the result simplifies to

$$E\{\eta_k \eta_m\} = \sigma^2 \mu_{km} \tag{7.39}$$

where

$$\mu_{km} = C_{km} \frac{\cos(\alpha_k - \alpha_m)}{a_k a_m} \tag{7.40}$$

This is the formulation we adopt for the Y_{ij} expression (6.10). Hence, applying (7.2), we get

$$E\{\zeta_k \zeta_m\} = [\sigma\lambda/(2\pi\hat{d})]^2 \mu_{km} \tag{7.41}$$

and the (previously) unknown factor of (6.10) is now expressible in terms of the μ_{km}'s as follows:

$$E\{\bar{\zeta}\bar{\zeta}\} = [\sigma\lambda/(2\pi\hat{d})]^2 \sum_{k,m=1}^{K} \mu_{km} \bar{e}_k \bar{e}_m \tag{7.42}$$

where K is the dimension of the output $\bar{e}$-space. Equivalently,

47

$$E\{\bar{\zeta}\bar{\zeta}\} = [\sigma\lambda/(2\pi\hat{d})]^2\, \bar{\bar{\mu}} \qquad\qquad (7.43)$$

where

$$\bar{\bar{\mu}} \equiv \sum_{k,m=1}^{K} \mu_{km}\bar{e}_k\bar{e}_m \qquad\qquad (7.44)$$

Equation (7.42) still contains one unknown, namely, σ^2. Let us elaborate: In principle, one could determine σ^2 by direct measurements. However, in practice, this raises a lot of problems such as calibration, variation of σ^2 with time and environmental conditions etc. Instead of that, one could obtain an estimate of σ^2 as an additional output of the least-squares processing. Specifically, with the proper processing, the vector of residuals $(\bar{v})$ yields an estimate of σ^2. Let us examine this processing.

The residuals vector $(\bar{v})$ is defined as follows:

$$\bar{v} \equiv \hat{\bar{p}} - \bar{\bar{A}}\cdot\hat{\bar{q}} \qquad\qquad (7.45)$$

Substituting for $\bar{\bar{A}}$ and $\hat{\bar{q}}$ from (5.6) and (5.8), respectively, we get

$$\bar{v} = \sum_{i=4}^{K} \bar{u}_i\bar{u}_i\cdot\hat{\bar{p}} \qquad\qquad (7.46)$$

But, for $i > 3$,

$$\bar{u}_i\cdot\bar{p} = 0 \qquad\qquad (i > 3) \qquad\qquad (7.47)$$

Hence (6.7),

$$\bar{u}_i\cdot\hat{\bar{p}} = \bar{u}_i\cdot\bar{\zeta} \qquad\qquad (i > 3) \qquad\qquad (7.48)$$

48

and (7.46) reduces to

$$\bar{v} = \sum_{i=4}^{K} \bar{u}_i \bar{u}_i \cdot \bar{\zeta} \qquad (7.49)$$

Using this form of $\bar{v}$, we find that $E\{v^2\}$ is expressible in terms of σ^2. Specifically,

$$E\{v^2\} = \sum_{i=4}^{K} \bar{u}_i \cdot E\{\bar{\zeta}\bar{\zeta}\} \cdot \bar{u}_i$$

$$= [\sigma\lambda/(2\pi\hat{d})]^2 \sum_{i=4}^{K} \bar{u}_i \cdot \bar{\bar{\mu}} \cdot \bar{u}_i \qquad (7.50)$$

In principle, (7.50) yields σ^2. In practice, we would like to avoid the computation of the (nonunique) set $\{\bar{u}_i\}_{i=4}^{K}$ This can be effected as follows:

$$\sum_{i=4}^{K} \bar{u}_i \cdot \bar{\bar{\mu}} \cdot \bar{u}_i = \left(\sum_{i=1}^{K} \bar{u}_i \cdot \bar{\bar{\mu}} \cdot \bar{u}_i \right) - \left(\sum_{i=1}^{3} \bar{u}_i \cdot \bar{\bar{\mu}} \cdot \bar{u}_i \right) \qquad (7.51)$$

Note that the first sum on the right is the trace of $\bar{\bar{\mu}}$ ($tr(\bar{\bar{\mu}})$). But the trace of a linear operator is an invariant which has the same value in any orthonormal frame. In particular,

$$tr(\bar{\bar{\mu}}) = \sum_{i=1}^{K} \bar{e}_i \cdot \bar{\bar{\mu}} \cdot \bar{e}_i = \sum_{i=1}^{K} \mu_{ii} \qquad (7.52)$$

This removes the last obstacle to the computation of σ^2 and yields:

$$[\sigma\lambda/(2\pi\hat{d})]^2 = \frac{E\{v^2\}}{\left(\sum_{i=1}^{K} \mu_{ii} \right) - \left(\sum_{i=1}^{3} \bar{u}_i \cdot \bar{\bar{\mu}} \cdot \bar{u}_i \right)} \qquad (7.53)$$

49

The covariance dyadic $E\{\bar{\zeta}\bar{\zeta}\}$ we set out to derive is thus finally given by

$$E\{\bar{\zeta}\bar{\zeta}\} = E\{v^2\} \frac{\bar{\bar{\mu}}}{\left(\sum_{i=1}^{K}\mu_{ii}\right) - \left(\sum_{i=1}^{3}\bar{u}_i \cdot \bar{\bar{\mu}} \cdot \bar{u}_i\right)} \tag{7.54}$$

where $\bar{\bar{\mu}}$ is given by (7.44) and (7.40).

Note that (7.40) is formulated in terms of the noiseless parameters a_k, α_k (see Fig. 13). Though these are not directly available to us, they are the statistical means of the corresponding noise-contaminated entities, which are available. One possible way of utilizing this fact in the application of (7.54) is as follows: We partition the input data into a large number M (say, $M = 100$) of independent data sets, apply the least-squares solution to each one of these sets, and obtain the DOA estimate from the mean of these M individual solution vectors $(\hat{\bar{q}})$. The covariance dyadic of this mean solution vector is just $1/M$ of the covariance dyadic $\bar{\bar{Y}}$ (6.3) of a single solution vector. Now, in evaluating $E\{\bar{\zeta}\bar{\zeta}\}$ required by the single-solution $\bar{\bar{Y}}$ (see (6.10)), we exploit the fact that we have at our disposal a large number (M) of independent sets of noisy a_k's and α_k's: We simply compute their means (over the M sets) as good estimates of the noise-free a_k's and α_k's required (indirectly) by (7.54) (see (7.40)). Similarly, for the factor $E\{v^2\}$ of (7.54), we substitute the mean of the M different values of v^2 obtained from the M least-squares solutions.

As we shall see in Section IX, this approach yields reasonable agreement with results based on direct covariance computations.

VIII. CONFIDENCE BOUNDS FOR THE DOA DETERMINATION (FINAL FORMULATION)

Having determined the noise covariance in the aperture plane, we resume now the development of the DOA confidence bounds. Our first step is to update the expression for the output covariance matrix (6.10) by substituting (7.42) in it:

$$Y_{ij} = \frac{[\sigma\lambda/(2\pi\hat{d})]^2}{z_i^2 z_j^2} \sum_{k,m=1}^{K} \mu_{km} (\bar{z}_i \cdot \bar{e}_k)(\bar{z}_j \cdot \bar{e}_m) \qquad (8.1)$$

All the variables appearing here are now available to us. Specifically,

$[\sigma\lambda/(2\pi\hat{d})]^2$	in equation (7.53)
z_i^2	in equations (5.26), (5.28), and (5.29)
K	in equation (3.27)
μ_{km}	in equations (7.40) and (7.38)
$\bar{z}_i \cdot \bar{e}_k$	in equations (5.15) through (5.17)

Thus, we have here all the information required to encode (in software) the computation of the output covariance matrix Y. This means that the central moments of the probability density function of the vector $\hat{\bar{q}}$ (Section VI) up to (and including) the second order, are now in our possession. Nothing has been said, however, about the higher moments of this density function. Yet, in order to obtain the desired DOA confidence bounds, we have to adopt a specific reasonable density function consistent with the computed central moments. Our choice is the three-dimensional Gaussian density function with the computed Y and $E\{\hat{\bar{q}}\}$ ($= \bar{q}$) as its parameters. Specifically, denoting

$$\bar{x} = \hat{\bar{q}} - \bar{q} = \sum_{i=1}^{3} x_i \bar{f}_i \qquad (8.2)$$

$$\bar{\bar{Z}} = \bar{\bar{Y}}^{-1} \quad (Z = Y^{-1}) \qquad (8.3)$$

we adopt the probability density [7]

51

$$G_3(\bar{x}) = \sqrt{\frac{|\bar{Z}|}{(2\pi)^3}} \; e^{-\frac{1}{2}\bar{x}\cdot\bar{\bar{z}}\cdot\bar{x}} \qquad\qquad (8.4)$$

Obviously, the surfaces of constant $G_3(\bar{x})$ in the three-dimensional $\bar{x}$-space are characterized by

$$\bar{x}\cdot\bar{\bar{Z}}\cdot\bar{x} = constant \qquad\qquad (8.5)$$

But this is an equation of an ellipsoid centered at the origin. This leads to ellipsoidal confidence bounds for the vector $\bar{q}$. However, whereas the first two components of $\bar{q}$ (n_1, n_2) prescribe the desired DOA, its third component $(a/\hat{d})$ is of no interest to us and unnecessarily complicates the interpretation of the results.

Such a situation is encountered in many other least-squares problems where it appears to be as unavoidable as it is in ours. The root of the problem seems to lie in the fact that, in some cases, the engineering parameters we try to extract are only part of a larger set of interrelated physical parameters. It is this interrelationship which dictates that the smaller subset of engineering parameters is retrievable only in the context of solving for the larger set of physical parameters. We claim, however, that though we cannot prevent the occurrence of such a scenario, we can cancel its undesirable effects. Let us examine how this is brought about.

$G_3(\bar{x}_0)$ provides the infinitesimal probability $dP_3(\bar{x}_0)$ of finding $\bar{x}$ in an infinitesimal (three-dimensional) region in $\bar{x}$-space around point $\bar{x}_0$, that is,

$$dP_3(\bar{x}_0) = G_3(\bar{x}_0) \; dx_1 \, dx_2 \, dx_3 \qquad\qquad (8.6)$$

However, we are interested in the probability of an (x_1, x_2) combination regardless of the value of x_3. To facilitate the formulation of such a probability, we introduce now

$$\hat{\bar{x}} = \sum_{i=1}^{2} x_i \, \bar{f}_i \qquad\qquad (8.7)$$

Thus,

$$dP_3(\bar{x}_0) = dP_3\left(\hat{\bar{x}}_0 + x_3 \, \bar{f}_3\right) \qquad\qquad (8.8)$$

and the probability that concerns us is

$$dP_2(\hat{\bar{x}}_0) = \int_{-\infty}^{\infty} dP_3(\hat{\bar{x}}_0 + x_3 \, \bar{f}_3) \, dx_3$$

$$= \left\{ \int_{-\infty}^{\infty} G_3(\hat{\bar{x}}_0 + x_3 \, \bar{f}_3) \, dx_3 \right\} \, dx_1 \, dx_2 \qquad (8.9)$$

This is equivalent to the introduction of a new two-dimensional density function

$$G_2(\hat{\bar{x}}) \equiv \int_{-\infty}^{\infty} G_3(\hat{\bar{x}} + x_3 \, \bar{f}_3) \, dx_3 \qquad (8.10)$$

which provides the infinitesimal probability $dP_2(\hat{\bar{x}}_0)$ of finding $\hat{\bar{x}}$ in an infinitesimal (two-dimensional) region in $\hat{\bar{x}}$-space around point $\hat{\bar{x}}_0$, that is,

$$dP_2(\hat{\bar{x}}_0) = G_2(\hat{\bar{x}}_0) \, dx_1 \, dx_2 \qquad (8.11)$$

With the adopted Gaussian distribution $G_3(\bar{x})$, the computation of $G_2(\hat{\bar{x}})$ is trivial: All we have to do is replace the three-dimensional covariance matrix Y by its upper left two-dimensional submatrix $\hat{Y}$ (see, for example, result 7.4.3 of [7] concerning the Gaussian marginal distribution). Thus, introducing the inverse

$$\hat{Z} = \hat{Y}^{-1} \qquad (8.12)$$

and regarding it as the representation of dyadic $\hat{\bar{\bar{Z}}}$ in the $\bar{f}_i$ frames, we get

$$G_2(\hat{\bar{x}}) = \sqrt{\frac{|\hat{Z}|}{(2\pi)^2}} \, e^{-\frac{1}{2} \hat{\bar{x}} \cdot \hat{\bar{\bar{Z}}} \cdot \hat{\bar{x}}} \qquad (8.13)$$

Having derived the explicit expression for $G_2(\hat{\bar{x}})$, we are now in position to compute the probability of finding $\hat{\bar{x}}$ inside any region in $\hat{\bar{x}}$-space by simply integrating (8.11) over that region. This computation becomes almost trivial when we select the $\hat{\bar{x}}$-space integration region to be an ellipse determined by the spectrum of $\hat{\bar{\bar{Z}}}$ and perform the integration via the corresponding elliptic annuli. Let us consider this in detail.

We start by tracing a curve in $\hat{\bar{x}}$-space along which $G_2(\hat{\bar{x}})$ has a constant value. Specifically, we choose the curve which satisfies

$$\frac{1}{2} \hat{\bar{x}} \cdot \hat{\bar{\bar{Z}}} \cdot \hat{\bar{x}} = t \tag{8.14}$$

where t is a fixed parameter. The shape of this curve is readily obtained from the spectrum of $\hat{\bar{\bar{Z}}}$ as follows: Note that the symmetry of $\bar{\bar{Y}}$ implies that $\hat{\bar{\bar{Z}}}$ is also symmetric. Applying the general SVD representation theorem (5.6) to the symmetric case, we see that $\hat{\bar{\bar{Z}}}$ is representable as

$$\hat{\bar{\bar{Z}}} = \sum_{i=1}^{2} \lambda_i \bar{v}_i \bar{v}_i \tag{8.15}$$

But, since the $\bar{v}_i$'s are orthonormal, this means that

$$\hat{\bar{\bar{Z}}} \cdot \bar{v}_i = \lambda_i \bar{v}_i \tag{8.16}$$

and we identify the λ_i's as the eigenvalues of $\hat{\bar{\bar{Z}}}$ and the $\bar{v}_i$'s as the corresponding eigenvectors.

Substitution of the (8.15) representation in (8.14) now yields

$$\frac{\xi_1^2}{2t/\lambda_1} + \frac{\xi_2^2}{2t/\lambda_2} = 1 \tag{8.17}$$

where

$$\xi_i = \hat{\bar{x}} \cdot \bar{v}_i \tag{8.18}$$

In other words, the selected curve is a member of a family of concentric ellipses with parameter t. All members of the family have their axes along the eigenvectors $\bar{v}_1$, $\bar{v}_2$ while their semiaxes depend on the parameter t as follows:

$$\text{length of } i\text{-th semiaxis} = \sqrt{2t/\lambda_i} \tag{8.19}$$

As we increase t, the size of the ellipse increases. Its orientation and shape, however, are not affected, being independent of t.

The area enclosed by any of these ellipses (A) is given by

$$A = 2\pi t / \sqrt{\lambda_1 \lambda_2} = 2\pi t / \sqrt{|\hat{z}|} \qquad (8.20)$$

so that the differential area of the infinitesimal annulus along the curve is given by

$$dA = \left(2\pi / \sqrt{|\hat{z}|}\right) dt \qquad (8.21)$$

Therefore (see (8.13) and (8.14)), the probability of finding $\hat{\bar{x}}$ in this annulus is ($e^{-t}dt$), and the probability of finding $\hat{\bar{x}}$ inside an ellipse with parameter t (denoted $P(t)$) is given by

$$P(t) = \int_0^t e^{-s}ds = 1 - e^{-t} \qquad (8.22)$$

The interpretation of this result is quite straightforward. Let us illustrate it with an example: Assume that we have applied the least-squares solution and our best estimate of the DOA is represented by point $\hat{D}$ on the unit sphere (see the first paragraph of Section VI). Now we want to find a 99% confidence region around this point. From (8.22) we find that the value of the corresponding t is 4.6. Substituting this value in (8.19), we find the semiaxes of the confidence ellipse, while its orientation is given by $\bar{v}_1$, $\bar{v}_2$, and its center is given by the least-squares solution. Strictly speaking, this ellipse is traced in the (x, y) plane of Fig. 1 and is the projection of the confidence region traced on the unit sphere. However since, in the usual circumstances, the center of this ellipse is very close to the origin and its size is very small compared to 1, this projected region is almost identical with the confidence region on the surface of the unit sphere.

IX.   SOFTWARE VERIFICATION

The main result established in the preceding sections is the development of two algorithms to process the outputs of the focal-plane horns. The first yields the best (least-squares) estimate of the DOA, while the second determines the confidence region around this estimate. Both algorithms have been implemented in software and are incorporated in a quite flexible program which provides easy selection of a large number of various operational parameters and allows the controlled injection of noise.

Though this program is capable of extensive testing of all aspects of the algorithms, so far we have carried out only the minimal number of tests sufficient to establish their validity. Since real data are not yet available, we have applied this program to simulated data provided to us by P.W. Cramer of JPL. These data refer to an idealized version (no deformation) of the 34-m antenna at DSS 13 of the Goldstone Deep Space Communications Complex. The assumed frequency is 32 GHz with a corresponding 3-dB loss at a deviation of about 8.4 millidegrees from boresight (the spherical coordinate θ) [8]. The data set consists of the outputs of conical horns of 4.4-in. (11.2-cm) diameter in configurations of up to three rings (see Fig. 2) and for incident waves arriving at zero to 12 millidegrees off boresight (in steps of 2 millidegrees) with several selected values of the spherical coordinate ϕ.

Our first concern was the behavior of the DOA algorithm in the absence of noise. Figure 14, which was obtained for the three-ring array $(N = 3)$ with $\phi = 0°$, is a reasonably good representative of the results obtained in all other cases $(N = 1,2; \phi = 90°,180°,270°)$. We see here a deviation of the estimated θ from the true θ which increases rapidly with the value of θ. We shall shortly introduce a plausible explanation for this behavior. However, it should be stressed at the outset that this deviation does not affect the viability of the algorithm. The basis for this statement is simply the fact that the observed deviation follows a well-defined, smooth curve and could thus be easily corrected via a "calibration curve" such as that of Fig. 14, regardless of the details of the error mechanism.

In searching for the possible cause of the deviation, we examined first the algorithm which translates the aperture-plane phase angles to the first two components of $\bar{n}$, the unit vector pointing to the source. This algorithm, which is given by equations (5.30), (5.31), and formulated in terms of templates in Figs. 11 and 12, is so simple that it can be easily tested with a hand calculator. Because of the linearity of the algorithm, it is sufficient to test it with just two properly chosen inputs. The inputs we have selected for this purpose are $\phi = 0°$ $(n_2 = 0)$ and $\phi = 90°$ $(n_1 = 0)$. According to (5.33), these correspond to

ESTIMATED θ (mdeg)
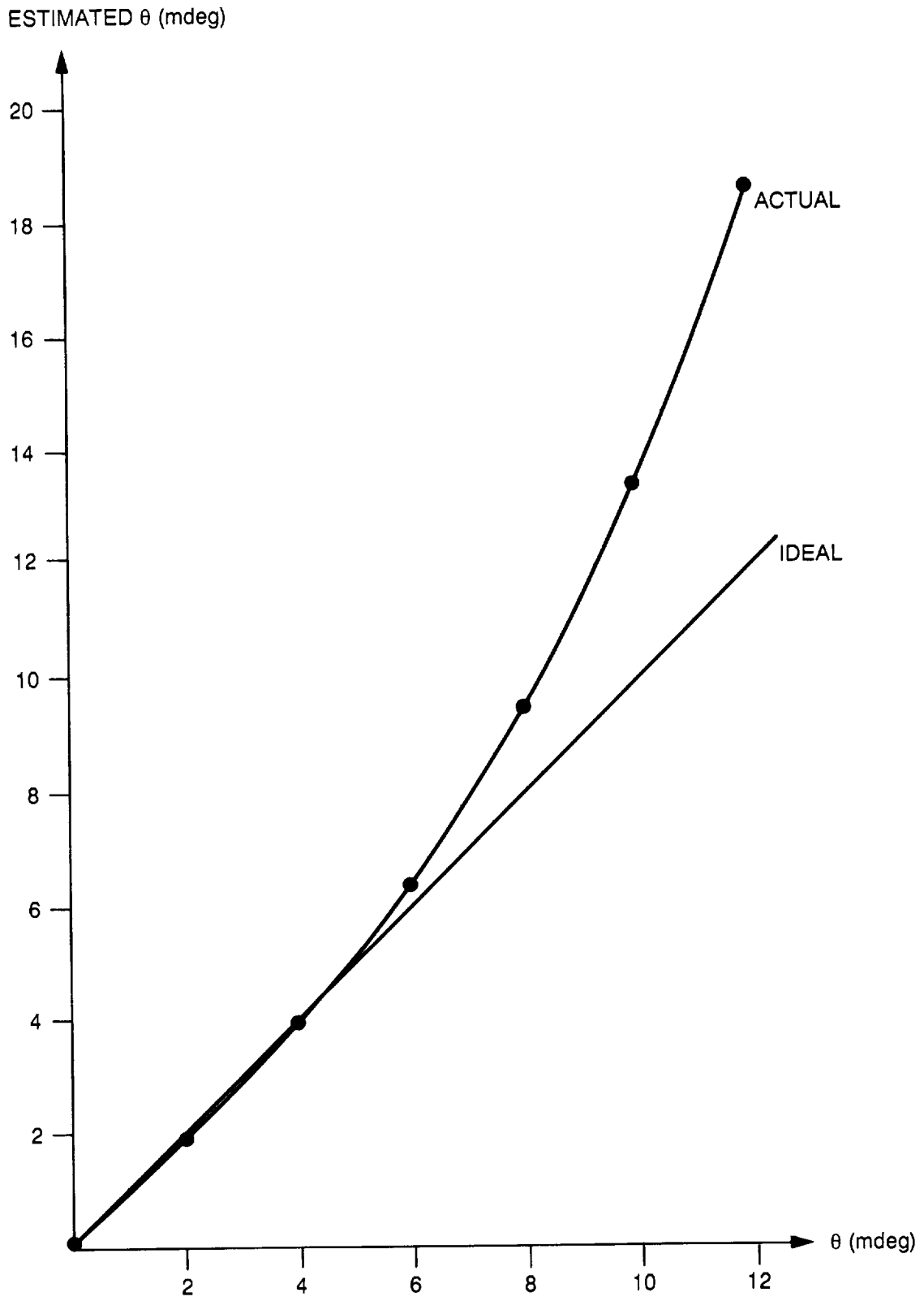

Fig. 14. θ "calibration curve" for the DOA algorithm $(N = 3)$.

$$n_1 = \sin \theta \quad (\phi = 0°) \qquad (9.1)$$

$$n_2 = \sin \theta \quad (\phi = 90°) \qquad (9.2)$$

Trivial trigonometry now shows that all values of $(w_k/\hat{d})$ are multiples of $\sin \theta$. Application of the templates of Figs. 11 and 12 to these computed values yields the correct results ((9.1) and (9.2)) thus verifying the validity of the algorithm.

We conclude that the problem of Fig. 14 must arise before the application of this algorithm, namely, in the transformation of the data from the focal plane to the aperture plane. In other words, the suspect is the Fresnel approximation and the thin-lens Fourier-transform property based on it (see Section III). Indeed, the very nature of the Fresnel approximation implies that it is applicable only to systems in which the angular deviation of rays from the optical axis (θ) is small. Now, the actual numerical value of θ which is considered small in this context, may differ quite a bit from system to system. One might argue, then, that Fig. 14 simply tells us that, for the DSS 13 antenna, the validity bound for the Fresnel approximation is about 5 millidegrees. The fact that the transition out of the Fresnel approximation region is as smooth as indicated in Fig. 14 is certainly a welcome result. As already pointed out, this allows us to extend the application of this approximation beyond this boundary.

It should be stressed that the above explanation is just a plausibility argument. The rigorous mathematical analysis required to substantiate it has not been carried out.

So far we have considered the behavior of the overall DOA algorithm with respect to the spherical coordinate θ. Its behavior with respect to the spherical coordinate ϕ is quite satisfactory. This fact lends further support to our explanation of the θ behavior.

We turn now to the second algorithm, namely, the one that determines the confidence region around the DOA estimate. Figure 15 is representative of what is obtainable with this algorithm. We show here a projection onto the (x, y) plane of a very small region of the unit sphere around the z axis. The tick marks along the axes correspond to the spherical coordinate θ (on the unit sphere) measured in millidegrees. Point D represents the true position of the unit vector pointing to the source $(\bar{n})$. Its spherical coordinates are $\theta = 0°.004$, $\phi = 0°$. Point $\hat{D}$ represents $\hat{\bar{n}}$ —the estimate of $\bar{n}$ under the following conditions:
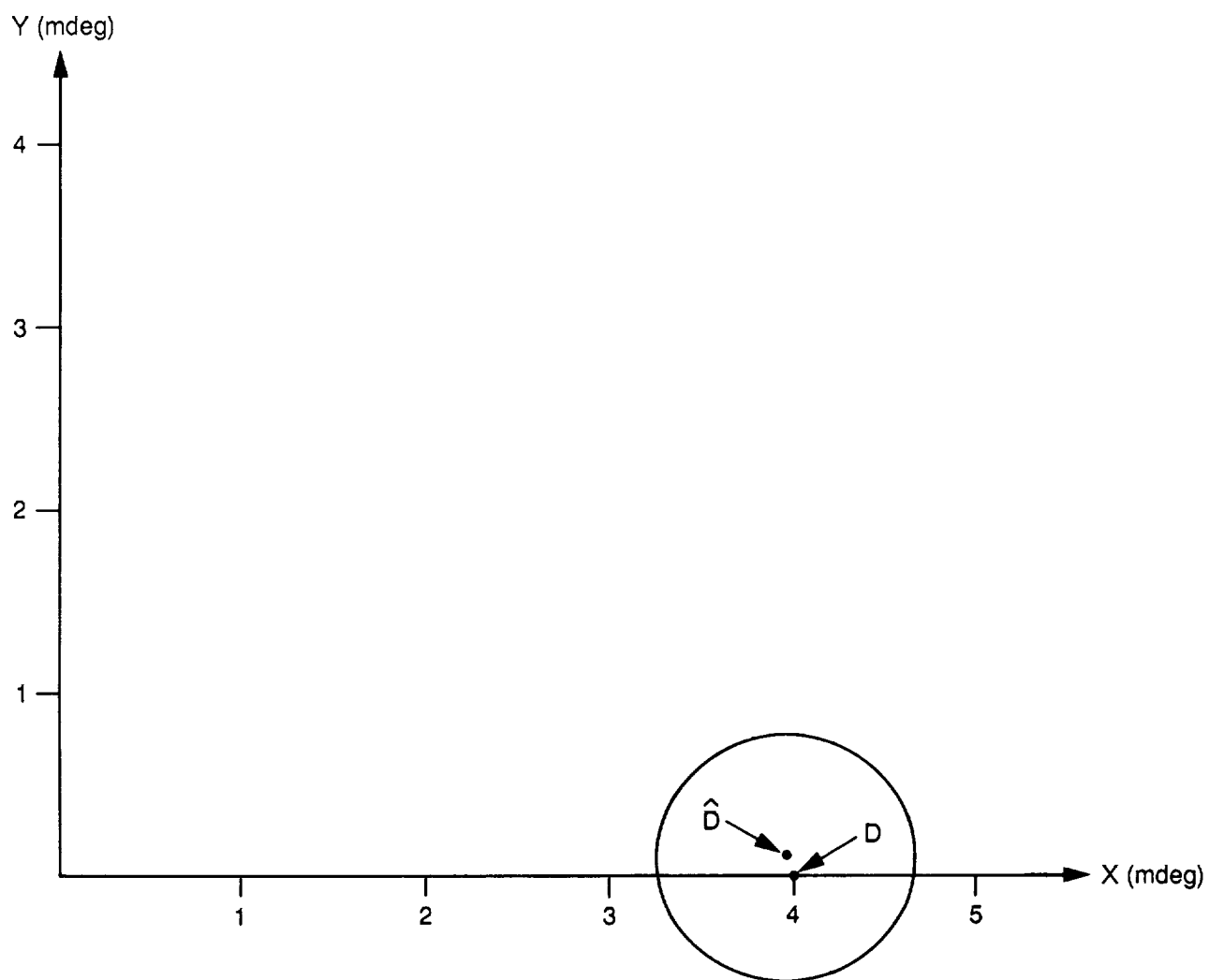
Fig. 15. Example of a theoretical confidence region $(N = 1)$.

$$\left.\begin{array}{l}\text{(signal power)/(noise spectral density)} = 40 \text{ dB-Hz} \\[4pt] \text{(number of samples used in the DOA estimate)} = 100 \\[4pt] \text{(integration time for each sample)} = 0.2 \text{ sec} \\[4pt] \text{(number of focal-plane rings, } N\text{)} = 1\end{array}\right\} \quad (9.3)$$

The ellipse shown centered at point $\hat{D}$ is the 99% confidence region computed by the second algorithm for this case; that is, it is the region around point $\hat{D}$ inside of which point D should be found with a probability of 99%. Note that point D, indeed, lies inside this ellipse. It should be pointed out that, though the region looks circular, it is in fact an ellipse with axes tilted with respect to the (x, y) axes.

To check the validity of this algorithm, we carried out a direct computation of the output covariance matrix Y (as opposed to its computation based on the theoretical (8.1)), using a much larger number of samples (2000). In Fig. 16, we show the confidence region corresponding to this direct computation superimposed on the theoretical confidence region of Fig. 15. Though the two are not identical, it is gratifying to note how close they are.

The practical operational significance of a result such as that in Fig. 15 is shown in Fig. 17 where we have added a circle of 3.15-millidegrees diameter to represent the locus of points where the loss would be 0.1 dB if the antenna is pointed at the estimated source direction $\hat{D}$ [8]. We see that the 99% confidence region lies well inside the 0.1-dB circle. Hence, there is at least a 99% probability that the loss due to the error in estimating $\bar{n}$ is less than 0.1 dB. Of course, when more detailed beam-pattern measurements become available, the single 0.1-dB circle would be replaced by a family of concentric circles for smaller losses, and a finer estimate of the probable bounds on the pointing loss would be obtained. In this case, one could determine a sequence of confidence regions corresponding to a sequence of signal integration times, thereby getting a clear presentation of the trade-off between signal integration time and pointing-error loss.

As we have already pointed out, Figs. 14 and 17 are typical of results obtained with other parameter combinations. Reviewing all these results, we conclude that both algorithms are viable.
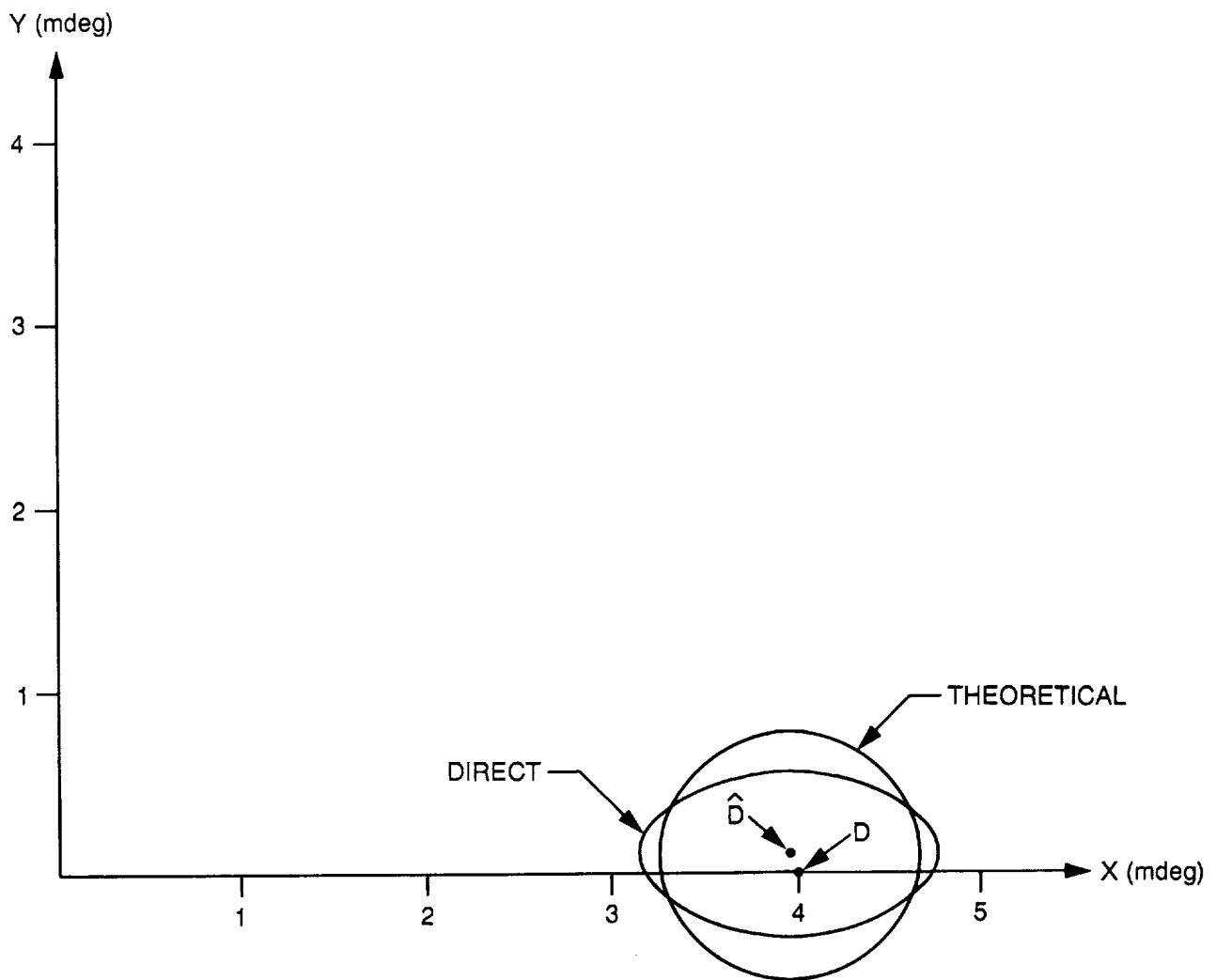
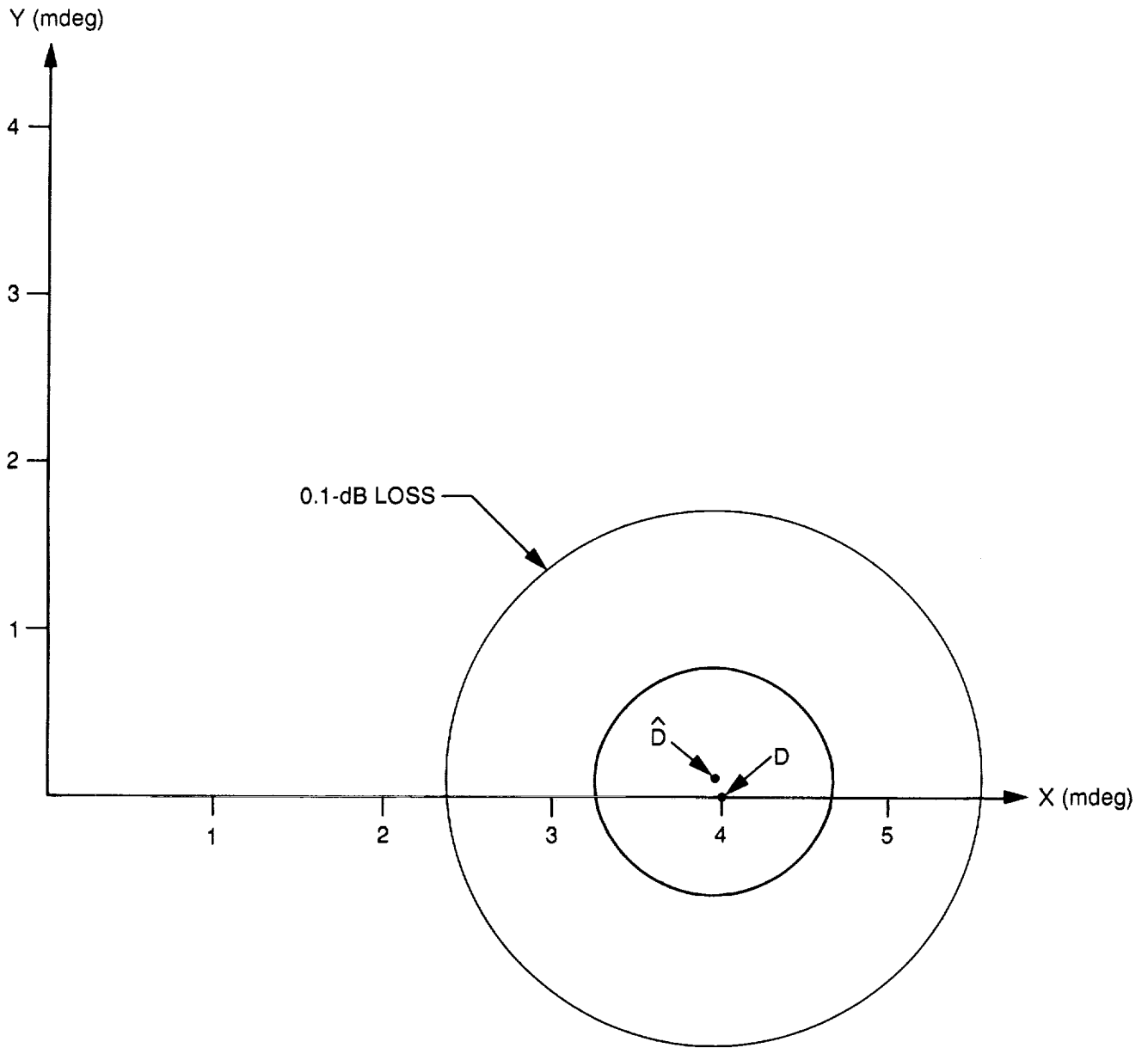Fig. 16. Comparison with a directly-computed confidence region.

Fig. 17. The pointing-error loss $(N = 1)$.

X.   CONCLUDING REMARKS

Consider the pointing of a highly directional antenna at a desired source. Assume now that the signal has been properly acquired so that, initially, there is no pointing error. As the source moves across the sky, the tracking system moves the antenna in an effort to keep the source close to the peak of the antenna pattern. A commonly used tracking system (CONSCAN) keeps the antenna in constant motion so that its axis (boresight) traces a small circle (on the unit sphere) centered at the nominal position of the source. When the actual position of the source drifts away from the center of the traced circle, the variation of the received power along the circle provides the information required to reposition the antenna and bring the source back to the center of the circle.

Note that with this scheme, the antenna is always pointing a little off the source, thus incurring a signal loss. To minimize this loss, the diameter of the scanned circle (expressed in degrees along a great circle on the unit sphere) should be much smaller than the 3-dB width of the antenna pattern. For highly directive antennas, this may lead to such a small circle that the mechanical part of the antenna-pointing system is incapable of tracing it with the required precision.

Against this background, consider now a tracking system based on the ideas developed here. We install in the focal plane of the antenna an array of horns (with their associated receivers) surrounding the customary single central collecting horn. Now, instead of constantly moving the antenna in a circular pattern around each point of its predicted tracking trajectory, we just process the horns' outputs, thereby getting a continuous monitoring of the pointing error. We issue a command to deviate from the predicted trajectory only in response to a sufficiently large detected pointing error. Nowhere in this system do we face the constraints of the mechanical pointing system.

Presumably, a pointing system using a focal-plane array has not been previously implemented because of the cost of the required additional hardware. As we have already pointed out, this argument does not apply to the case of our immediate concern because the hardware referred to is an integral part of the deformation-compensation system. However, when a very large single-dish antenna is considered, the investment in a focal-plane array would probably be only a small fraction of the total cost of the antenna. In this case, then, the antenna-pointing system proposed here would be a viable approach even for a nondeformed antenna.

We have shown that the DOA algorithm developed here is equivalent to the solution of the purely geometrical problem of finding the best (least-squares) estimate of a plane from the noisy measurements of its distances from a large set of given points. In Appendix B we came across a different purely geometrical equivalent of the DOA problem. This alternative formulation just replaces the plane of the adopted formulation with a point. Specifically, we found in Appendix B that the DOA problem is also equivalent to finding the best (least-squares) estimate of a point from the

noisy measurements of its distances from a large set of given points. Though we have not followed through on this alternative approach, it seems worthwhile to examine the algorithm it generates and compare it to the one developed here.

Another subject which merits further study is the effect of the finite areas of the focal-plane horns. Our analysis regards the horns as idealized point-samplers of the focal-plane field. In other words, we assume that the voltage output of a horn is proportional to the field at the center of its aperture. In reality, this voltage is proportional to the integral of the field over the aperture. Thus, our assumption is equivalent to approximating the field at the center of the horn aperture by the average of the field over the full horn aperture. For sufficiently small horns, this is a reasonable approximation. Though the simulation results of Section IX seem to provide an indirect validation of this assumption, the subject merits closer scrutiny.

Finally, it would be very worthwhile to study the derived output-covariance expression (8.1) and apply the insight gained to the selection of optimal system parameters. In this context, it should be pointed out that the adopted set of aperture-plane points of (3.34) underlying (8.1) is just one of several alternative candidates. Optimization analysis would probably prove fruitful here too.

APPENDIX A: REAL DYADICS

A dyad $(\bar{\bar{D}})$ is an entity represented by two vectors as follows (recall that we use overbars to denote vectors):

$$\bar{\bar{D}} = \bar{a}\bar{b}$$

(A-1)

There is no implied operation between $\bar{a}$ and $\bar{b}$. The meaning of the dyad is defined in terms of what it does to other vectors. Thus,

$$\bar{\bar{D}} \cdot \bar{h} = (\bar{a}\bar{b}) \cdot \bar{h} = \bar{a}(\bar{b} \cdot \bar{h})$$

(A-2)

that is, the result of (scalarly) postmultiplying $\bar{\bar{D}}$ with $\bar{h}$ is the vector $\bar{a}$ multiplied by the scalar $(\bar{b} \cdot \bar{h})$. Note that, in general, the vectors $\bar{a}$, $\bar{b}$ would belong to different spaces and would have different dimensionalities. The only constraint in (A-2) is the obvious one that $\bar{h}$ and $\bar{b}$ belong to the same space.

Equation (A-2) shows $\bar{\bar{D}}$ to be a (rather limited) linear operator operating on vectors in the $\bar{b}$-space to yield vectors in the $\bar{a}$-space. This same dyad can also operate in the opposite direction. Thus,

$$\bar{y} \cdot \bar{\bar{D}} = \bar{y} \cdot (\bar{a}\bar{b}) = (\bar{y} \cdot \bar{a})\bar{b}$$

(A-3)

Alternatively, this "opposite-direction" operation may be implemented using the concept of transposition. Specifically, the transpose of the dyad $\bar{\bar{D}}$ (denoted $\tilde{\tilde{D}}$) is given by

$$\tilde{\tilde{D}} = \bar{b}\bar{a}$$

(A-4)

Hence, (A-3) may be expressed, equivalently, as

$$\bar{\bar{D}} \cdot \bar{y} = (\bar{b}\bar{a}) \cdot \bar{y} = \bar{b}(\bar{a} \cdot \bar{y}) \qquad \text{(A-5)}$$

Though a dyad is a very limited linear operator, a linear combination of dyads (called a dyadic) can represent any linear operator. Specifically, given the set of dyads $\{\bar{\bar{H}}_i\}_{i=1}^m$, we can construct the dyadic $\bar{\bar{H}}$ as a linear combination of its elements as follows:

$$\bar{\bar{H}} = \sum_{i=1}^m \gamma_i \bar{\bar{H}}_i \qquad \text{(A-6)}$$

where the γ_i's are scalars.

It should be pointed out that the above dyadic $\bar{\bar{H}}$ may also be represented as a different linear combination of a different set of dyads. In other words, representation (A-6) is not unique. There is, however, a unique canonical representation of dyadics which is very useful and provides a powerful tool in manipulating linear operators in general. Specifically, an arbitrary dyadic (linear operator) $\bar{\bar{A}}$ of rank k is always representable as follows:

$$\bar{\bar{A}} = \sum_{i=1}^k s_i \bar{u}_i \bar{v}_i \qquad \text{(A-7)}$$

where

$\{\bar{v}_i\}_{i=1}^k$ is an orthonormal set spanning the domain of $\bar{\bar{A}}$ (A-8)

$\{\bar{u}_i\}_{i=1}^k$ is an orthonormal set spanning the range of $\bar{\bar{A}}$ (A-9)

$\{s_i\}_{i=1}^k$ is a set of positive scalars (the singular values) (A-10)

This is the SVD (Singular Value Decomposition) representation.

The special dyadic

$$\bar{\bar{I}} = \sum_{i=1}^{n} \bar{g}_i \, \bar{g}_i \qquad \text{(A-11)}$$

constructed from an orthonormal set $\{\bar{g}_i\}$ spanning the n-dimensional space is the identity dyadic in that space. Indeed, representing $\bar{a}$ in that frame as

$$\bar{a} = \sum_{i=1}^{n} \alpha_i \, \bar{g}_i \qquad \text{(A-12)}$$

we get

$$\bar{\bar{I}} \cdot \bar{a} = \left(\sum_{i=1}^{n} \bar{g}_i \bar{g}_i \right) \cdot \left(\sum_{j=1}^{n} \alpha_j \, \bar{g}_j \right)$$

$$= \sum_{i,j=1}^{n} \bar{g}_i \delta_{ij} \, \alpha_j = \sum_{j=1}^{n} \alpha_j \bar{g}_j = \bar{a} \qquad \text{(A-13)}$$

Similarly,

$$\bar{a} \cdot \bar{\bar{I}} = \bar{a} \qquad \text{(A-14)}$$

The properties established so far allow us to prove statement (2.5), namely, that the gradient of a position vector is the identity dyadic. To do that, we let the vector $\bar{a}$ be a position vector in the n-dimensional space. In this case, the gradient operator $\bar{\nabla}$ is defined in terms of the components of $\bar{a}$ as

$$\bar{\nabla} \equiv \sum_{i=1}^{n} \bar{g}_i \frac{\partial}{\partial \alpha_i} \qquad \text{(A-15)}$$

67

Hence,

$$\bar{\nabla}\bar{a} = \left(\sum_{i=1}^{n} \bar{g}_i \frac{\partial}{\partial\alpha_i}\right)\left(\sum_{j=1}^{n} \bar{g}_j \alpha_j\right) \qquad \text{(A-16)}$$

But

$$\frac{\partial\alpha_j}{\partial\alpha_i} = \delta_{ij} \qquad \text{(A-17)}$$

Hence,

$$\bar{\nabla}\bar{a} = \sum_{i,j=1}^{n} \delta_{ij} \bar{g}_i \bar{g}_j = \sum_{i=1}^{n} \bar{g}_i \bar{g}_i = \bar{\bar{I}} \qquad \text{(A-18)}$$

and (2.5) is proved.

The scalar product of two dyads is another dyad. Let

$$\bar{\bar{F}} = \bar{c}\bar{d} \qquad \text{(A-19)}$$

then

$$\begin{aligned}
\bar{\bar{D}} \cdot \bar{\bar{F}} &= (\bar{a}\bar{b}) \cdot (\bar{c}\bar{d}) \\
&= \bar{a}(\bar{b} \cdot \bar{c})\bar{d} \\
&= (\bar{b} \cdot \bar{c})(\bar{a}\bar{d}) \qquad \text{(A-20)}
\end{aligned}$$

Similarly, the product of two dyadics is another dyadic.

Finally, the inverse of a nonsingular dyadic $\bar{\bar{N}}$, denoted $\bar{\bar{N}}^{-1}$, satisfies

$$\bar{\bar{N}}^{-1} \cdot \bar{\bar{N}} = \bar{\bar{I}} \qquad (A-21)$$

$$\bar{\bar{N}} \cdot \bar{\bar{N}}^{-1} = \bar{\bar{I}} \qquad (A-22)$$

We conclude with a derivation of the relationship between the dyadic representation of the operation of a linear operator and the more familiar matrix representation. Let the dyadic $\bar{\bar{A}}$ operate on the vector $\bar{q}$ in the n-dimensional input space (spanned by the orthonormal set $\{\bar{f}_j\}_{j=1}^{n}$) to yield the vector $\bar{p}$ in the m-dimensional output space (spanned by the orthonormal set $\{\bar{e}_i\}_{i=1}^{m}$). The dyadic representation of this process is

$$\bar{\bar{A}} \cdot \bar{q} = \bar{p} \qquad (A-23)$$

Now we replace $\bar{q}$ by $\bar{\bar{I}} \cdot \bar{q}$, where $\bar{\bar{I}}$ is the identity operator of the input space, and dot both sides of (A-23) by $\bar{e}_i$:

$$\bar{e}_i \cdot \bar{\bar{A}} \cdot \left(\sum_{j=1}^{n} \bar{f}_j \, \bar{f}_j \right) \cdot \bar{q} = \bar{e}_i \cdot \bar{p} \qquad (A-24)$$

Extending the j-summation to cover all of the left side, we get

$$\sum_{j=1}^{n} (\bar{e}_i \cdot \bar{\bar{A}} \cdot \bar{f}_j)(\bar{f}_j \cdot \bar{q}) = (\bar{e}_i \cdot \bar{p}) \qquad (A-25)$$

But this has essentially the matrix form we are after. Specifically, denoting

$$(\bar{e}_i \cdot \bar{p}) = p_i \; = \; element \; i \; of \; the \; column \; matrix$$
$$representing \; \bar{p} \qquad (A-26)$$

69

$$(\bar{f}_j \cdot \bar{q}) = q_j = \textit{element j of the column matrix}$$

$$\textit{representing } \bar{q} \tag{A-27}$$

$$(\bar{e}_i \cdot \bar{\bar{A}} \cdot \bar{f}_j) = A_{ij} = \textit{element (i, j) of the rectangular matrix}$$

$$\textit{representing } \bar{\bar{A}} \tag{A-28}$$

we get the standard matrix representation of process (A-23) in the adopted orthonormal frames, namely,

$$\sum_{j=1}^{n} A_{ij} q_j = p_i \tag{A-29}$$

(A-28) also provides the foundation for converting in the other direction. Let us use it to substitute for A_{ij} in the following double summation:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \bar{e}_i A_{ij} \bar{f}_j = \left(\sum_{i=1}^{m} \bar{e}_i \bar{e}_i \right) \cdot \bar{\bar{A}} \cdot \left(\sum_{j=1}^{n} \bar{f}_j \bar{f}_j \right) \tag{A-30}$$

We have here two identity operators: the right one operates on the input space while the left one operates on the output space. In view of (A-7), (A-13), and (A-14), both of them may be dropped and we get

$$\bar{\bar{A}} = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} \bar{e}_i \bar{f}_j \tag{A-31}$$

Note that, whereas the (A-31) representation of $\bar{\bar{A}}$ is a sum of mn dyads (corresponding to the mn elements of the matrix A), its SVD representation is a sum of only k dyads (where k is the rank of $\bar{\bar{A}}$). This, however, requires knowledge of its singular value decomposition.

A good treatment of the subject of dyadics is given in [9, chapter IV]. A more extensive treatment combined with the SVD formulation is given in [4].

APPENDIX B: THE FOCAL-PLANE DISTRIBUTION DUE TO AN OFF-BORESIGHT
PLANE WAVE

In Section III we dealt with the determination of the aperture-plane distribution from a given (measured) focal-plane distribution. Here we consider the reverse problem, namely, finding the focal-plane distribution due to an aperture-plane distribution generated by a (slightly) off-boresight incident plane wave. This is motivated by the need to find the rate of attenuation of the focal-plane field as we move away from the optical axis. This rate affects the selection of N —the number of rings in Fig. 2.

We start with the aperture-plane distribution $\hat{u}$ caused by a plane wave with propagation vector $\bar{k}$ (see (2.1),(3.9))

$$\hat{u} = e^{i(\bar{k}\cdot\hat{\bar{r}}-\omega t)} \tag{B-1}$$

Expressing this in terms of the normalized aperture-plane position vector $\hat{\bar{R}}$ (see (3.6),(3.8)), we get

$$\hat{u}(\hat{\bar{R}}) = e^{i(L\bar{k}\cdot\hat{\bar{R}}-\omega t)} \tag{B-2}$$

Now we apply (3.13) to get the corresponding focal-plane distribution $u(\bar{R})$, namely,

$$u = e^{i(\alpha+\pi R^2-\omega t)} \iint_{\hat{S}} e^{i(L\bar{k}\cdot\hat{\bar{R}}-2\pi\bar{R}\cdot\hat{\bar{R}})} d\hat{S} \tag{B-3}$$

where $\hat{S}$ refers to the aperture area and $d\hat{S}$ is an element of this area which we now express explicitly in its polar form

$$d\hat{S} = (\hat{r}d\psi)\, d\hat{r} = L^2\hat{R}\, d\hat{R}\, d\psi \tag{B-4}$$

Hence,

$$U \equiv u e^{-i(\alpha+\pi R^2-\omega t)} = L^2 \int_{\hat{R}=0}^{\hat{R}_a} \int_{\psi=0}^{2\pi} e^{i(L\bar{k}-2\pi\bar{R})\cdot\hat{\bar{R}}} \hat{R}\, d\hat{R}\, d\psi \tag{B-5}$$

where $\hat{R}_a$ is the normalized radius of the antenna aperture.

71

Next we turn our attention to the exponent in the integrand of (B-5). Substituting for $\bar{k}$ using (2.3),(2.16), and noting that $\hat{\bar{R}} \cdot \bar{h}_3 = 0$, we find that the exponent takes the following form:

$$-i2\pi \left[(L/\lambda)\bar{n}_s + \bar{R} \right] \cdot \hat{\bar{R}} \qquad (B-6)$$

This suggests the introduction of the "offset position vector" $\bar{b}$ given by

$$\bar{b} = F\bar{n}_s + \bar{r} \qquad (B-7)$$

where F is the focal length of the antenna and $\bar{r}$ is the (unnormalized) position vector in the focal plane (3.10). The exponent (B-6) is now expressible in terms of $\bar{B}$, the normalized version of $\bar{b}$, given by

$$\bar{B} = \bar{b}/L = (L/\lambda)\bar{n}_s + \bar{R} \qquad (B-8)$$

This simplifies (B-5), yielding

$$U = L^2 \int_{\hat{R}=0}^{\hat{R}_a} \int_{\psi=0}^{2\pi} e^{-i2\pi\bar{B}\cdot\hat{\bar{R}}} \; \hat{R} \; d\hat{R} \; d\psi \qquad (B-9)$$

In evaluating this integral, we measure the polar angle ψ in reference to the vector $\bar{B}$ as shown in Fig. 18. This means that

$$\bar{B} \cdot \hat{\bar{R}} = B\hat{R} \cos \psi \qquad (B-10)$$

and (B-9) now takes the following form:

$$U = L^2 \int_{\hat{R}=0}^{\hat{R}_a} \left[\int_{\psi=0}^{2\pi} e^{-i2\pi B\hat{R}\cos\psi} \; d\psi \right] \hat{R} \; d\hat{R} \qquad (B-11)$$
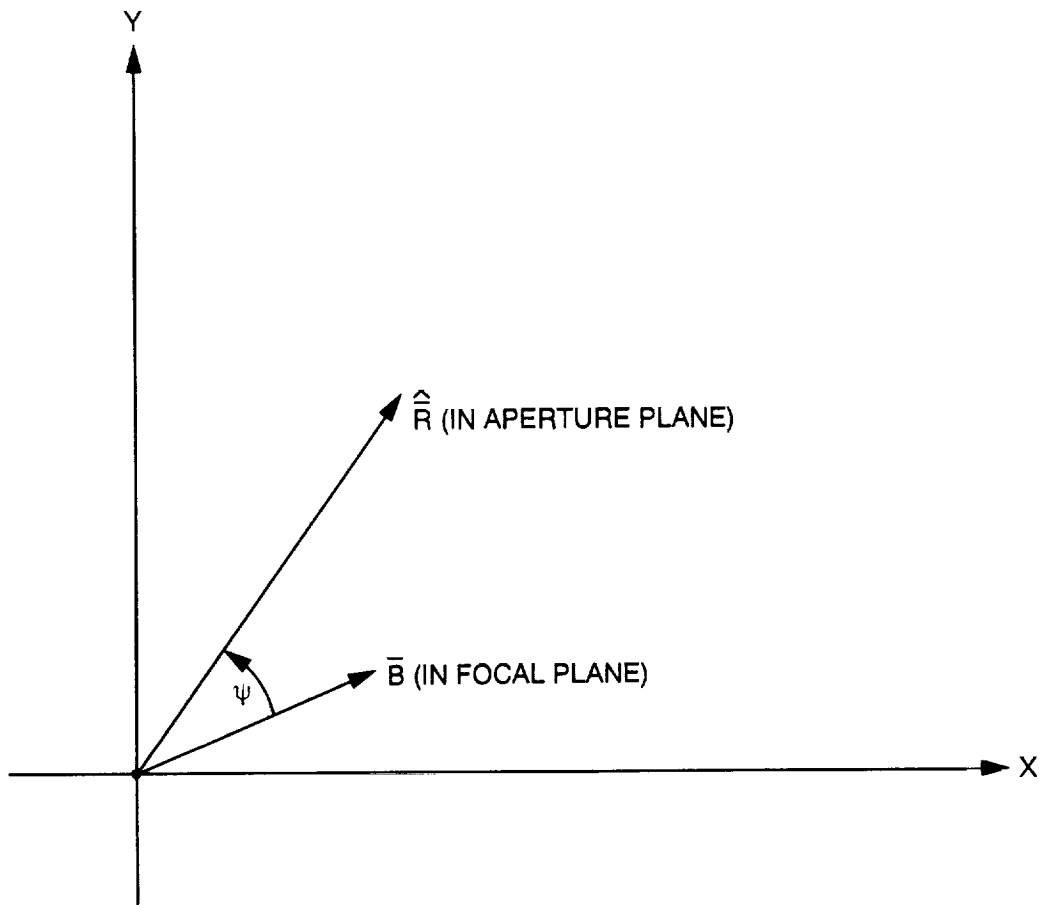
Fig. 18. Polar coordinates in the aperture plane.

The bracketed integral is now recognized as a multiple of the Bessel function of order zero (J_o) (see formula 9.1.21 of [10]). Specifically,

$$U = 2\pi L^2 \int_{\hat{R}=0}^{\hat{R}_a} J_o(2\pi B\hat{R}) \; \hat{R} \, d\hat{R}$$

$$= [(L/B)^2/(2\pi)] \int_{0}^{2\pi B\hat{R}_a} x J_0(x) \; dx \qquad \text{(B-12)}$$

But the integral appearing here is immediate and given by $xJ_1(x)$ where $J_1(x)$ is the Bessel function of order 1 (see formula 9.1.30 of [10]). This yields

$$U = U(B) = \left(2\pi \hat{r}_a^2\right) \frac{J_1(2\pi B\hat{R}_a)}{(2\pi B\hat{R}_a)} \qquad \text{(B-13)}$$

where $\hat{r}_a$ is the unnormalized radius of the antenna aperture.

Note that for small x,

$$J_1(x) \approx (1/2)x \quad (|x| \ll 1) \qquad \text{(B-14)}$$

(see formula 9.1.7 of [10]). This implies

$$\lim_{x \to 0} J_1(x)/x = 1/2 \qquad \text{(B-15)}$$

Furthermore, when we superpose the line $y = (1/2)x$ on a plot of $y = J_1(x)$ (such as Fig. 9.1 of [10]) we find that

$$|J_1(x)| \le |(1/2)x| \qquad \text{(B-16)}$$

We conclude that the function $J_1(x)/x$ as well as its magnitude assume their maximal values of $1/2$ at the origin. Applying this to (B-13), we get

$$|U(B)| \le U(0) = \pi \hat{r}_a^2 = \text{area of aperture} \qquad \text{(B-17)}$$

Hence the normalized phrasing of (B-13):

$$\frac{U(B)}{U(0)} = 2 \frac{J_1(2\pi B\hat{R}_a)}{(2\pi B\hat{R}_a)} \qquad \text{(B-18)}$$

Note that U differs from the field entity u only in phase (see (B-5)). Hence the maximal value of $|u|$ occurs at $\bar{B} = \bar{0}$. Thus, the coordinate of the peak response $\bar{r}_{PEAK}$ is given by

$$\bar{r}_{PEAK} = -F\bar{n}_s \tag{B-19}$$

(see (B-7)) while the corresponding normalized coordinate $(\bar{R}_{PEAK})$ is given by

$$\bar{R}_{PEAK} = -(L/\lambda)\bar{n}_s \tag{B-20}$$

(see (B-8)). It follows that the value of u at this point (u_{PEAK}) is given by

$$u_{PEAK} = U(0)\, e^{i\{\alpha + \pi[(L/\lambda)n_s]^2 - \omega t\}} \tag{B-21}$$

and finally

$$\frac{u}{u_{PEAK}} = 2\,\frac{J_1(2\pi B\hat{R}_a)}{(2\pi B\hat{R}_a)}\, e^{i\pi\{R^2 - [(L/\lambda)n_s]^2\}} \tag{B-22}$$

Note that the vectors $\bar{R}$ and $\bar{B}$ affect this expression only via their magnitudes. This means that the structure of the focal-plane field is describable in terms of two families of concentric circles (see Fig. 19); the constant-magnitude loci, which are circles of radius B centered at $\bar{R}_{PEAK}$ and the constant-phase loci, which are circles of radius R centered at the origin.

Applying formula 9.2.1 of [10] to (B-22), we find that for large values of the argument $(2\pi B\hat{R}_a)$,

$$|u/u_{PEAK}| < \sqrt{8/\pi}\Big/(2\pi B\hat{R}_a)^{\frac{3}{2}} \qquad (2\pi B\hat{R}_a \gg 1) \tag{B-23}$$

For the parameters of Fig. 7, this yields for the fourth ring

$$|u/u_{PEAK}| < 0.007 \qquad (N = 4) \tag{B-24}$$

and thereby we are justified in ignoring terms of the fourth and higher rings. Note, however, that even when we use only one ring, the neglected terms are still quite small. Thus,

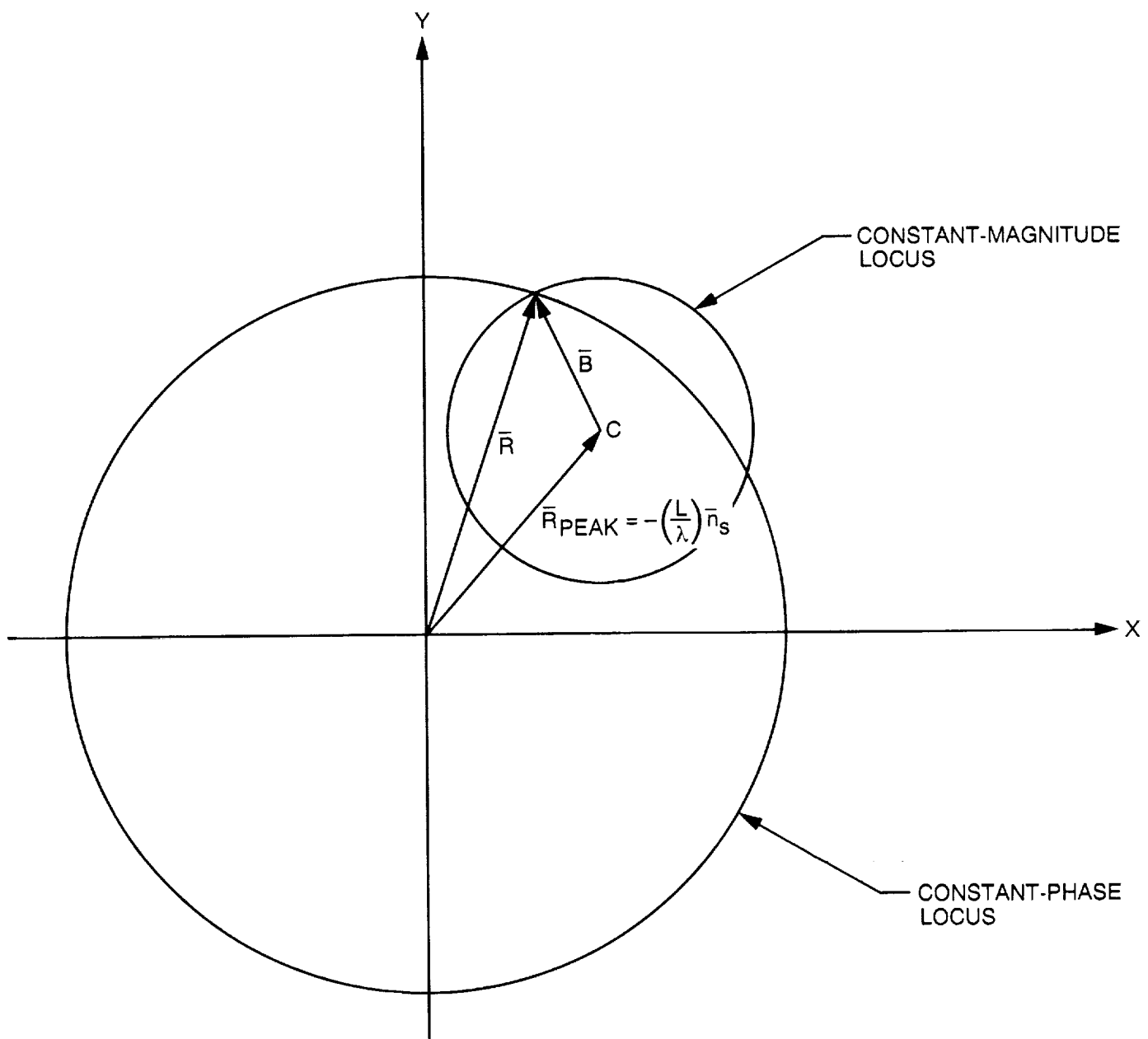$$|u/u_{PEAK}| < 0.021 \qquad (N = 2) \tag{B-25}$$

Fig. 19. Structure of focal-plane field.

Before leaving this subject, it should be pointed out that (B-22) and Fig. 19 provide (at least in principle) an alternative approach to the DOA extraction. For example, if we subtract from the phase of each (u/u_{PEAK}) the known phase πR^2 and average these entities over all horns, we get an estimate of n_s^2 and hence the θ coordinate of the source. Alternatively, one could ignore the phase information and extract the parameter $(2\pi B\hat{R}_a)$ from each focal-plane voltage. Each B is the (normalized) distance from the known location of the measurement to the tip of the vector $\bar{R}_{PEAK}$ (point C in Fig. 19). Obviously, if we succeed in locating point C, we have obtained $\bar{R}_{PEAK}$ and, hence, $\bar{n}_s$ (B-20), and the problem is solved. It is interesting to note that, as in the adopted formulation of Section IV, we have here a purely geometric phrasing of the DOA extraction, namely, finding a point (C) from its distances (the B's) to a given set of points. Here, we are dealing with distances to a point. There, we have been dealing with distances to a plane.

An optimal combination of this approach with the one yielding n_s^2 is worth looking into. Whether the resulting algorithm would be practically feasible and advantageous is not currently known.

REFERENCES

[1]   V.A. Vilnrotter, E.R. Rodemich, and S.J. Dolinar, Jr., "Real-Time Combining of Residual Carrier Array Signals Using ML Weight Estimates," *IEEE Trans. Comm.*, vol. 40, no. 3, pp. 604–615, March 1992.

[2]   J.W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 1968. New York.

[3]   C.L. Lawson and R.J. Hanson, *Solving Least Squares Problems*, Prentice-Hall, 1974. Englewood Cliffs, New Jersey.

[4]   S. Zohar, *The Modus Operandi of Real Linear Operators*, JPL internal report JPL D-9096 47 pp., Jet Propulsion Laboratory, October 1991. Pasadena, California.

[5]   L.B.W. Jolley, *Summation of Series*, second revised edition, Dover, 1961. New York.

[6]   V.A. Vilnrotter, "Channel Assignments and Array Gain Bounds for the Ka-Band Array Feed Compensation System," *TDA Progress Report 42-106*, vol. April-June, Jet Propulsion Laboratory, August 15, 1991. Pasadena, California.

[7]   S.S. Wilks, *Mathematical Statistics*, John Wiley & Sons, 1962. London.

[8]   P. Cramer, "DSS 13 Predict Study for Horn at F1 Focus," JPL internal document, IOM 3328-90-0345, Jet Propulsion Laboratory, August 29, 1990. Pasadena, California.

[9]   L. Brand, *Vector and Tensor Analysis*, John Wiley & Sons, 1947. London.

[10]   M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Applied Mathematics Series 55, National Bureau of Standards. Washington, D.C.