Radio Science, Volume 32, Number 6, Pages 2373–2391, November–December 1997

Coherent radar imaging: Signal processing and statistical properties

Ronald F. Woodman
Jicamarca Radio Observatory, Instituto Geofísico del Perú, Lima

Abstract. The recently developed technique for imaging radar scattering irregularities has opened a great scientific potential for ionospheric and atmospheric coherent radars. These images are obtained by processing the diffraction pattern of the backscattered electromagnetic field at a finite number of sampling points on the ground. In this paper, we review the mathematical relationship between the statistical covariance of these samples, $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^\dagger \rangle$, and that of the radiating object field to be imaged, $\langle \mathbf{F}\mathbf{F}^\dagger \rangle$, in a self-contained and comprehensive way. It is shown that these matrices are related in a linear way by $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^\dagger \rangle = \mathbf{a}\mathbf{M}\langle \mathbf{F}\mathbf{F}^\dagger \rangle \mathbf{M}^\dagger \mathbf{a}^*$, where $\mathbf{M}$ is a discrete Fourier transform operator and $\mathbf{a}$ is a matrix operator representing the discrete and limited sampling of the field. The image, or brightness distribution, is the diagonal of $\langle \mathbf{F}\mathbf{F}^\dagger \rangle$. The equation can be linearly inverted only in special cases. In most cases, inversion algorithms which make use of a priori information or maximum entropy constraints must be used. A naive (biased) "image" can be estimated in a manner analogous to an optical camera by simply applying an inverse DFT operator to the sampled field $\hat{\mathbf{f}}$ and evaluating the average power of the elements of the resulting vector $\hat{\mathbf{F}}$. Such a transformation can be obtained either digitally or in an analog way. For the latter we can use a Butler matrix consisting of properly interconnected transmission lines. The case of radar targets in the near field is included as a new contribution. This case involves an additional matrix operator $\mathbf{b}$, which is an analog of an optical lens used to compensate for the curvature of the phase fronts of the backscattered field. This "focusing" can be done after the statistics have been obtained. The formalism is derived for brightness distributions representing total powers. However, the derived expressions have been extended to include "color" images for each of the frequency components of the sampled time series. The frequency filtering is achieved by estimating spectra and cross spectra of the sample time series, in lieu of the power and cross correlations used in the derivation.

1. Introduction

One of the most important advances in the last few years in the field of atmospheric and ionospheric radar techniques is the use of imaging techniques. The first successful observations were obtained at the Jicamarca Radio Observatory, Lima, Peru, using the "coherent" radar mode [*Kudeki and Sürücü*, 1991; *Sürücü*, 1992; *Hysell*, 1996; *Hysell and Woodman*, this issue].

Copyright 1997 by the American Geophysical Union.

Paper number 97RS02017.
0048-6604/97/97RS-02017$11.00

A coherent radar makes use of enhanced index of refraction fluctuations produced by either ionospheric or atmospheric irregularities. Wind profilers, mesosphere-stratosphere-troposphere radars, and radars for the study of E and F region ionospheric irregularities are examples of such radars. Imaging the irregularities which produce the scattering is providing a better understanding of the physical mechanisms responsible for their formation.

Some of the radar "images" of irregularities previously shown in the literature [e.g., *Woodman and La Hoz*, 1976], although loosely interpreted as two-dimensional (east-west and vertical) pictures, are actually range-time-intensity (RTI) plots. RTI plots could be interpreted as slit-camera pictures, but with their inherent limitations. The limitations come

about because of the nonfrozen nature of the target and its possible two-dimensional horizontal structure in the most general case. In this paper, we consider real three-dimensional images in time. The time dimension is represented as a time sequence of images in a movie fashion. Although three-dimensional images (horizontal cross section and range) have not been attempted yet, we include their theory in this review. It just involves a redefinition of the indices in our vector and matrix notation.

Doppler sorting the echoes adds one further dimension. It makes possible three-dimensional "color" movies. The number of colors can be as large as desired, depending on the structure of the Doppler-frequency spectrum. The study of F region irregularities, with their large number of spectral features [*Woodman and La Hoz,* 1976], can benefit greatly from this possibility. In general, the possibility of imaging in coherent radars opens a whole new horizon in the field of atmospheric and ionospheric research.

Imaging techniques have been used in radar and radio astronomy for quite some time. There is extensive literature on the subject [e.g., *van Schooneveld,* 1979a; *Ostro,* 1993, and references therein]. Ionospheric and atmospheric radar imaging techniques can benefit from this experience, and indeed they have done so. Nevertheless, there are sufficient differences with these astronomical applications to warrant making a review of the fundamentals of coherent radar imaging and its signal processing requirements in a single comprehensive unit. This is the goal of this paper. It is hoped that it will be useful to students and researchers entering this new field.

Radar astronomy takes advantage of the rotation of the planets and of radar ranging to produce an image. Radio astronomy has a frozen sky to image (and plenty of time to do so) and can take advantage of the rotation of the Earth to change the projected baseline of an interferometer pair. The different baselines are used to synthesize a large aperture with different antenna spacings. Neither of these possibilities is available for coherent radars. Furthermore, the ionospheric and atmospheric irregularities are always changing, and their depth requires multiple images, which have to be obtained in a relative short time. Since there are requirements specific to coherent radar imaging, which the astronomical systems do not have, and vice versa, a specialized treatment is justified.

Several new contributions that are included in this review are the following: (1) the focusing of images in the near field; (2) a discussion of the statistical nature of the processes involved; (3) the introduction of a vector and matrix notation, which simplifies the treatment of two-dimensional images (actually three, when range is included) and their full covariance statistics; and (4) the representation of sampling and aperture truncation as a linear matrix operator.

There is a closer similarity between a coherent radar image system and an optical camera than with the other radio astronomy counterparts. We have taken advantage of this similarity in what follows in order to provide a better physical insight to their common mathematical models. Coherent radar images change so quickly that the similarity is closer to a movie camera than to a still one. In fact, most of the images obtained so far have been put in the form of video images (E. Kudeki and D. Hysell, personal communications, 1996). Although it is not possible to publish them in present journals, they can be and are distributed via the World Wide Web. These "movies" are of great importance to the understanding of the physical dynamical processes they depict.

A great difference between a coherent radar and a camera is the ability to sample and process the actual radio electromagnetic field (not just the power) in a digital form. In the radar case, the field can be sampled and processed at the proper space and time Nyquist rate, while it is impossible to do the same in optics with present technology. Thus the analogy can be used only to get physical insight into the mathematical algorithms, not for the processing techniques. Even in the radar case, the computational demand is very great, especially when one considers that there are multiple images corresponding to the different radar ranges. Fortunately, the information content per unit time is not as large as in optics, and computers are becoming faster and very economical. It is not a coincidence, then, that coherent radar imaging has been achieved only recently.

It is expected that the trend in rapid computer development will continue. Therefore we have reviewed the processing needs and algorithms, ignoring the possible practical processing limitations with current technology, and considered applications that have not been implemented so far. Along these lines, we present the theory of two-dimensional imaging with arbitrary two-dimensional antenna distributions. We will consider special simpler cases, amenable to implementation with present technology, as particu-

lar ones derived from the more general approach. One particular case discussed is the use of a one-dimensional nonredundant spacing array, such as the one used recently at Jicamarca [*Kudeki, and Sürücü*, 1991; *Hysell*, 1996; *Hysell and Woodman*, this issue].

We describe the mathematical model of the process using matrix notation. Once the symbols are introduced, this notation allows us to present the mathematics with greater simplicity. The fields are represented as column vectors. Matrix operators, with a single-letter symbol, represent the field propagators, the effect of an aperture, and that of a focusing lens. The operators act on the field column vector by matrix inner products. The same operator, and its transpose, operates on the full covariance statistical matrices of these column vectors to obtain the statistics of other linearly related fields.

Evaluating the statistics of the full covariance matrix of the observable field presents no additional complexity. It allows us to discuss the effect of some linear operations that spoil the diagonal character of the covariance matrix of the initial radiation field that we are trying to image. Having expressions for the evaluation of the covariance matrices is the first step in making an error analysis of the estimated image. In addition, optimum inversion algorithms require full statistical description of the processes involved. Simpler statistical concepts, like correlation functions and power spectra, are obtained from the principal (elements with the same index) and secondary (transverse) diagonals of these matrices.

The statistical mathematical problem is posed as one of modeling the system, from reality to image estimate, and then as one of inversion. In certain instances, it is possible to invert by simply applying matrix inverse operations. In others, it is not. Unfortunately, some of the matrices are singular, requiring modeling techniques with a priori constraints, including those imposed by maximum entropy algorithms. We do not cover the latter. We stop at posing the inversion problem and specifying the constraints. For its actual inversion, we refer the reader to the extensive literature on maximum entropy methods [e.g., *Smith and Grandy*, 1985; *Skilling and Sibisi*, 1996] and to the recent results presented by *Hysell* [1996].

An additional difference between an imaging radar and an optical camera is produced by the ranging possibilities of the first. An optical image projects all the objects between the camera and infinity onto a single plane. In a radar, it is possible to discriminate between different surfaces by delay ranging to produce multiple separate images, one for each different range. In our discussions, we will assume that this is always possible but will discuss only the imaging processes corresponding to a single range.

We will first discuss the imaging of the total power backscattered by the target. However, as in the optical case, it is possible to have as many images as "colors" (Doppler shifted spectral ranges) one wishes to filter out. In the last section, we discuss what is involved in producing multiple color images as a modification to the earlier discussions.

2. The Mathematical Model of Imaging

In order to gain some physical insight into the definitions of some of the fields and their relation to imaging, we will first consider the mathematical model of an optical camera and draw an analogy to a radar imaging system. We will define some of the notation in the process.

In an optical camera the lenses have two functions: first, to collimate the rays coming from an object in the near (Fresnel) field before they go through the aperture and second, to focus them again in the focal plane. These two functions are usually obtained with two lenses (or sets of lenses), one above and the other below the iris (aperture). The need for the second function never exists in the radar case; one can always project the image at infinity (Fraunhofer field). The radar analog for the first lens is necessary for images of "objects" in the near field. However, we will discuss first, at length, radar imaging of objects in the far field, which is the most common case. Thus we require no lenses. In optics this is the case of a pinhole camera. At first, then, we need to consider the simple mathematical model of a pinhole camera with no lenses. In a later section, we will treat the case of radar near-field imaging and the implementation of a radar lens analogy.

2.1. A Pinhole Camera: A Fourier Transform Operator

Figure 1a depicts schematically a pinhole camera (the lens shown is needed for later discussion). There are three important planes: the object plane, the aperture plane, and the focal, or image, plane. The size of the pinhole or aperture in the aperture plane has been exaggerated for illustration purposes. The following statements and approximations will be valid only if the distances between planes, L_1 and L_2,

are much larger than the pinhole. The condition to meet is that the two planes, object and focal, have to be in the far field of the aperture. That is, their distances to aperture ratio, L_1/a and L_2/a, have to be much larger than the aperture measured in wavelengths, a/λ. The pinhole can be, and in fact normally is, very large in terms of a wavelength.

At any one instant, the optical electromagnetic field at the three planes can be expressed as a complex function of the respective two-dimensional positions, $\boldsymbol{\theta}, \mathbf{x}$, and $\widehat{\boldsymbol{\theta}}$. We shall use only their complex scalar amplitudes, ignoring the vectorial nature of the fields, to keep the mathematical symbolism simple. There are four fields in these three planes: $F(\boldsymbol{\theta})$ and $\widehat{F}(\widehat{\boldsymbol{\theta}})$, corresponding to the object and focal plane, and $f(\mathbf{x})$ and $\widehat{f}(\mathbf{x})$, corresponding to the field just above and below the aperture plane. The function $\widehat{f}(\mathbf{x})$ is equal to $f(\mathbf{x})$, for $\mathbf{x}$ within the aperture, and zero elsewhere. It is convenient to represent the object and focal plane field functions in terms of an angular set of coordinates. We will take them with respect to a set of orthogonal coordinates aligned with the axis of the camera and centered at the center of the aperture. We shall use the directional cosines of a unit vector corresponding to the point in question. Thus we define $\boldsymbol{\theta} = [\sin(\theta_x), \sin(\theta_y)]$. Similarly, $\widehat{\boldsymbol{\theta}} = [\sin(\widehat{\theta}_{\mathbf{x}}), \sin(\widehat{\theta}_{\mathbf{y}})]$, but with the negative direction of the z axis. It is also convenient to express the $\mathbf{x}$ coordinates in wavelength units, i.e., $\mathbf{x} = [kx_1, kx_2]$. In terms of these coordinates, it can be shown [e.g., *Ratcliffe*, 1956] that

$$F(\boldsymbol{\theta}) \rightleftharpoons cf(\mathbf{x}) \tag{1}$$

$$\widehat{F}(\widehat{\boldsymbol{\theta}}) \rightleftharpoons \widehat{c}\widehat{f}(\mathbf{x}), \tag{2}$$

where $\rightleftharpoons$ stands for a Fourier transform pair with a kernel of the form $\exp(\pm i\boldsymbol{\theta}\cdot\mathbf{x})$. Note that we do not need to make the usual approximation $\boldsymbol{\theta} \approx [\theta_x, \theta_y]$, although its validity occurs so often that it justifies our notation: the use of $\boldsymbol{\theta}$ for the set of directional cosines.

Notice we have used the same symbol for the corresponding functions above and below the aperture plane, except for a modifying hat. A symbol with a hat implies that the corresponding function has been modified somehow by the aperture truncation. We have used this notation convention throughout the paper. In addition, we have used $\widehat{\boldsymbol{\theta}}$ for the angular variable of the image. Although it has the same magnitude as the angle of the corresponding object point, $\boldsymbol{\theta}$, the hat reminds us of its independence and that the angle is measured with respect to the same camera axis but in the opposite direction to the one used for $\boldsymbol{\theta}$.

If we ignore for a while, as a zero-order approximation, the effect of the aperture truncation and assume $\widehat{f}(\mathbf{x}) = f(\mathbf{x})$, the above equations tell us that there is a one-to-one mapping; that $\widehat{F}(\widehat{\boldsymbol{\theta}}) = CF(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$. In other words, within a constant of proportionality, C, the field at the focal plane is a homologous image of the object. We could have reached this conclusion using arguments borrowed from ray optics, but then we could not have evaluated the real effect of truncation. Let us see next what the effect of truncation is.

Let us introduce a function $a(\mathbf{x})$ to represent the effect of the aperture such that

$$a(\mathbf{x}) = 1, \mathbf{x} \text{ within the aperture}$$

$$a(\mathbf{x}) = 0 \text{ elsewhere,} \tag{3}$$

and such that

$$\widehat{f}(\mathbf{x}) = a(\mathbf{x})f(\mathbf{x}), \tag{4}$$

and define

$$A(\boldsymbol{\theta}) \rightleftharpoons a(\mathbf{x}), \tag{5}$$

where $\rightleftharpoons$ stands for a Fourier transform pair with the same kernel as before.

It can be shown, using the convolution theorem, that the relationship between the object and the focal fields is then

$$\widehat{F}(\widehat{\boldsymbol{\theta}}) = CA(\boldsymbol{\theta}) * F(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}. \tag{6}$$

The asterisk above implies a convolution operation. The imaging therefore still exists, but $\widehat{F}(\widehat{\boldsymbol{\theta}})$ is a blurred or smeared representation of $F(\boldsymbol{\theta})$. The smearing (convolving) function $A(\boldsymbol{\theta})$ is a narrow function of $\boldsymbol{\theta}$ of the order of $\lambda/2a$, where a is the scale size of $a(\mathbf{x})$. This operation is illustrated in Figure 1a by showing less structure in $\widehat{F}(\widehat{\boldsymbol{\theta}})$ than in $F(\boldsymbol{\theta})$.

Note again that a pinhole, despite its name, cannot be modeled or approximated by $a(\mathbf{x}) = \delta(\mathbf{x})$. Its finite size, measured in wavelengths, is important. In fact, the larger the aperture the better the performance of the pinhole camera (the same for the radio analog). As we have mentioned above, it needs to be small only with respect to the distances to the object and the image plane, L_1 and L_2.

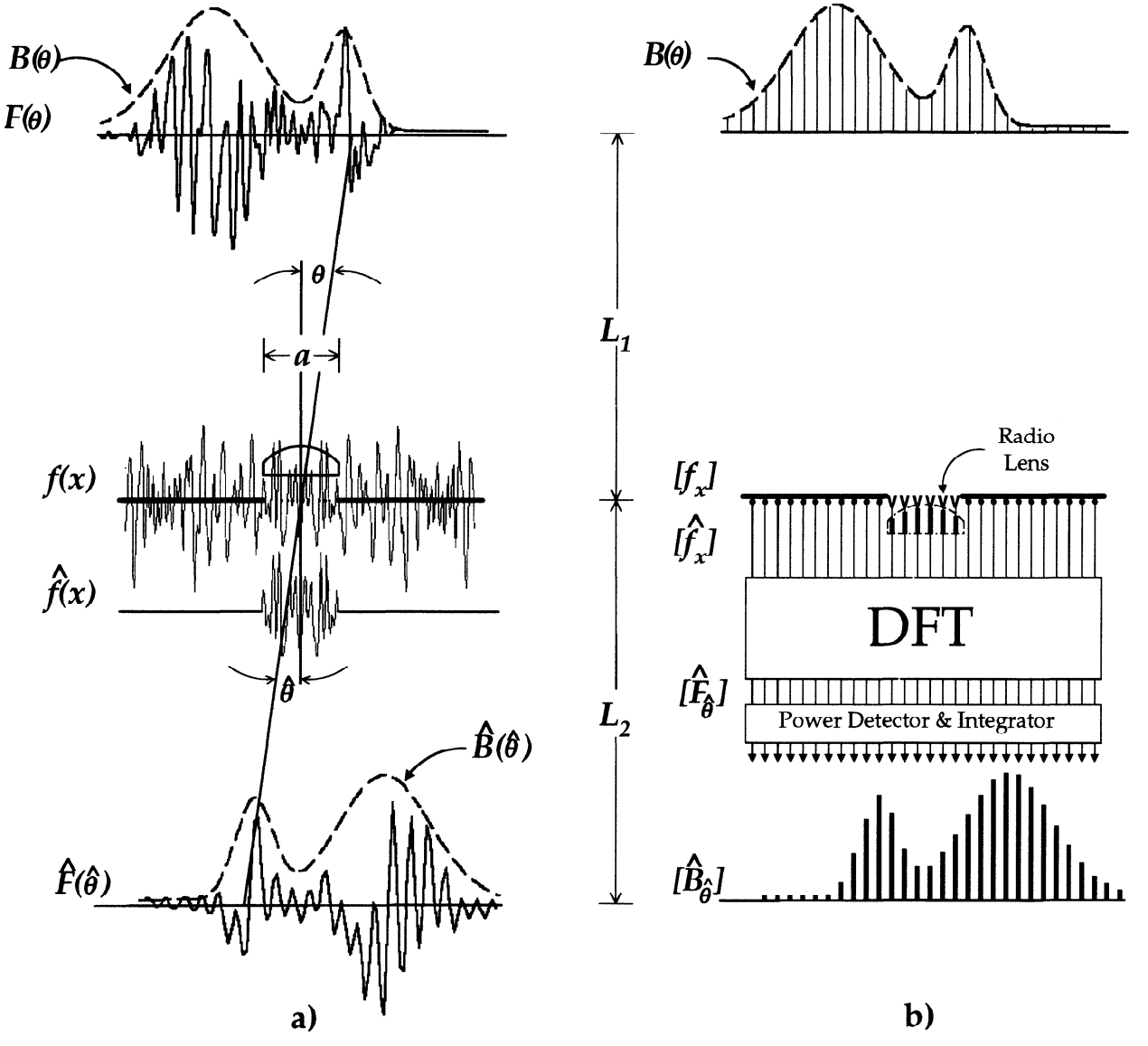

Figure 1. (a) Schematic representation of a pinhole camera. (b) Schematic realization of a discrete "radio camera" by means of an antenna array and a DFT "black box" drawn as an overlay of Figure 1a. Refer to the main text for symbol definitions. The lenses are needed only for objects in the near field.

We derive three important conclusions from the above discussion: (1) A pinhole camera is a perfect Fourier transform operator acting on the field present within the aperture (equation (2)); (2) all the information about the image (not the object!) is present in the field at the aperture; and (3), the image is a smeared representation of the object (even under proper focus, as we shall see later). The same holds true at radio frequencies, but obviously it is impossible to construct a pinhole radio camera; the dimensions would be enormous if we are to satisfy the $L_2/a >> a/\lambda >> 1$ condition.

The analogy is only physically realizable down to the aperture. In order to obtain an image, we need to find alternative ways of performing the same Fourier transformation performed by the propagation within

the space between the aperture and the image plane represented by equation (2). A digital approach is obvious, but later, and in the appendix, we will also discuss one possible analog implementation. In either case, we would have to sample the field at the aperture plane. For this we need an array of antennas within the aperture, i.e., where $a(\mathbf{x}) \neq 0$. To stress the optical analogy, we could even imagine making a hole in the opaque ground (not necessary!) and placing an even larger array of sampling antennas slightly under, including the region outside the aperture. Only the ones inside the aperture would sample nontrivial values of the field, so we can obviate the construction of the ones below the opaque ground and replace them and their corresponding receivers with dummy lines with zero amplitudes. Such a conceptual picture is depicted in Figure 1b. The analog to the camera proper would then be a "black box" (of realizable dimensions), as the one depicted in Figure 1b, capable of performing the Fourier transformation of this sampled and truncated $\widehat{f}(\mathbf{x})$ field, $[\widehat{f}_{\mathbf{x}}]$. The inputs to the box are the output of the antennas and corresponding receivers plus the mentioned dummy lines, and the outputs would represent the sampled image field, $[\widehat{F}_{\boldsymbol{\theta}}]$. With this picture in mind, we can identify the corresponding radio fields $f(\mathbf{x})$ and $\widehat{f}(\mathbf{x})$ as the fields just above and below the hypothetical perforated opaque ground. The set $[f_{\mathbf{x}}]$ and $[\widehat{f}_{\mathbf{x}}]$ stand for the sampled representations of the same fields. Notice that we have used brackets and indices for the discrete representations of the corresponding continuous functions. The symbol for the index is the same as the one used for the independent variable in the corresponding continuous function. We shall use this notation convention throughout the paper.

There is an additional operation performed in a camera that needs to be discussed. It is the detection and recording of the image. In an optical photographic camera, this is performed by a film, sensitive to the average intensity (integrated power) of the field at the focal plane, $\widehat{B}(\widehat{\boldsymbol{\theta}}) = \langle \widehat{F}(\widehat{\boldsymbol{\theta}}) \widehat{F}(\widehat{\boldsymbol{\theta}})^* \rangle$. Notice we have used a hat over B to differentiate it from the real brightness of the object, $B(\boldsymbol{\theta}) = \langle F(\boldsymbol{\theta}) F(\boldsymbol{\theta})^* \rangle$, it is trying to image. This operation can be performed in the radio case by standard analog, digital, or hybrid techniques, all of which involve the averaging of a detected power at each point of the equivalent "focal plane", that is, at each output of our Fourier transform black box, as illustrated in Figure 1b. This reduces, on the one hand, the amount of required information to be recorded, and on the other, it requires a statistical treatment of the instrument, given the random nature of the fields. The recorded information is depicted as sampled lines in Figure 1b labeled $[\widehat{B}_{\boldsymbol{\theta}}]$.

At this point, it is worth mentioning that the usual case of a radar or radio antenna connected to a single receiver and power detector is analogous to a camera with a single grain of light-sensitive recording emulsion at only one point in the focal plane. This is certainly a waste of information. What we try to achieve with imaging is the full use of a given aperture (antenna) to detect in parallel the radiation intensity coming from several different directions, i.e., to make use of all the information available in the electromagnetic field available at the aperture of the antenna, just as the optical camera does.

We should also mention that in photographic optics (excluding holography) one is seldom interested in the actual instantaneous electromagnetic fields $F(\boldsymbol{\theta})$ and $\widehat{F}(\widehat{\boldsymbol{\theta}})$. In fact, they are never available as observables. A description of the brightness of the object $B(\boldsymbol{\theta})$ and its image $\widehat{B}(\widehat{\boldsymbol{\theta}})$ is sufficient. However, this is not the case in radio imaging. In radio and radar imaging, one does sample the instantaneous electromagnetic fields. This justifies our concern with the properties of the former.

We are confronted with an additional conceptual problem. $F(\boldsymbol{\theta})$ can be highly structured. Its structural size is determined by the coherence length of the scatterers, (which can be as small as a Debye length in the case of an incoherent scatter radar) and by their distance from the aperture defining the angle they subtend. If, as an approximation, we place the object field at infinity, the angular size of the structure would reduce to an infinitesimal width. This very small structure, in turn, because of its Fourier relationship with $f(\mathbf{x})$, implies the need for a very large domain for $\mathbf{x}$. Fortunately, this large domain is only a mathematical construct, which can be as large as necessary. The practical implications will be discussed later in the image estimation section. At this time it is sufficient to say that it can be taken as finite with no practical consequences.

2.2. Discretization of the Aperture and Image Plane Field

The optical camera performs a continuous transformation on continuous fields. In drawing the radar analogy, we have already mentioned the need for a

discrete sampling of the fields. Therefore it is appropriate that we discuss its effects.

The presence of $\sin(\theta)$ terms in the kernel of the Fourier transform automatically limits the maximum spatial frequency in the aperture field. The maximum values that the integration variables, $\sin(\theta_x)$ and $\sin(\theta y)$, take are limited by $|\boldsymbol{\theta}| = (\sin^2\theta_x + \sin^2\theta_y)^{1/2} = 1$. Larger values correspond to evanescent waves, which are not in the radiation fields considered here. Thus $F(\theta)$ is "band-limited" and $f(\mathbf{x})$ can be sampled at the Nyquist "frequency", which can be shown to be $\lambda/2$. Furthermore, in accordance with the previous section the $\mathbf{x}$ domain can be taken as finite, and therefore $f(\mathbf{x})$ can be represented by a finite number N of samples, $f_{\mathbf{x}}$. It is also up to us to define the shape of the domain. For convenience in evaluating its discrete Fourier transform, we will define it as being rectangular and much larger than the aperture. Within this domain we will take $N = n \times m$ evenly distributed samples. The aperture, on the other hand, can have an arbitrary shape. We will use the same domain, for the time being below the aperture plane, and represent $\widehat{f}(\mathbf{x})$ by the same N number of samples, $\widehat{f}_{\mathbf{x}}$. Of course, as we mentioned before, we do not need antennas or receivers to sample the zero values of the field outside the aperture. Furthermore, we will later discuss computation economical ways of reducing N, taking advantage of the trivial values outside the aperture. This finite set of samples, $[\widehat{f}_{\mathbf{x}}]$, contains all the information available at the aperture and, consequently, its discrete Fourier transform (DFT), $[\widehat{F}_{\widehat{\boldsymbol{\theta}}}]$, contains the same information.

The sampling of the field can be obtained by short dipoles, $\lambda/2$ apart. However, we are not forced to do this if we are only interested in obtaining images in a limited region of the sky. Just as in the case of a time series, we can filter the $[f_{\mathbf{x}}]$ sequence and sample it at a new and "slower" Nyquist rate (decimation). The filter equivalent in our case is defined by the physical size of the antenna sampling element (not to be confused with the size of the imaging array), which can be larger than $\lambda/2$. To each sampling antenna element we can associate a weighting function confined within its physical size, for example, 1 within the antenna element and 0 elsewhere. The output at the feed end of this element is then the convolution of the radio wave field with the weighting function of the antenna, sampled at the position of the electrical center of the antenna. The effect of this filtering in the aperture plane is to reduce the angular portion of the sky (object plane) to an angular area defined by the pattern of the antenna sampling element, in the same way as time filtering reduces the width of the frequency content of a time series. An aperture filled with these sampling antenna elements, spaced at distances comparable to the size of the antenna elements (proper Nyquist separation), does not lose any further information. The new system would be capable of producing an image within the pattern of these larger antenna elements and with a resolution defined by the even larger array defining the aperture.

There is no direct optical analog to this spatial filtering at the aperture plane. The closest would be to block waves arriving from outside a central region with a black cone (lens hood) extending from the pinhole or by sensing a limited area in the focal plane. The effect is similar but it is achieved in a different way.

In the radar case, there is an additional way of "filtering" and selecting a given small region of the sky. One can use a transmitter antenna that only illuminates this given region. In what follows, we will assume we have done so and that the proper sampling antenna size and separation has been used. Furthermore, we will assume that the combined transmitter antenna and sampling receiver antenna element weigh to zero the object field outside the main beam of the sampling antennas. In reality, it is impossible to achieve this in an absolute way. In practice, it is important to keep this limitation in mind, since there is an implicit periodicity in the DFT which would always alias undesired regions of the weakly illuminated object outside the region into the main beam of the sampling antenna .

2.3. Imaging in Matrix Notation

Being able to discretize continuous object, aperture, and image field functions into a set of N samples allows us to use simple matrix notation. The notation allows us to model mathematically a discrete imaging system regardless of the method, analog or digital, used to perform the DFTs. The sampled field functions, introduced before, can be represented by the following column "vectors":

$$\widehat{\mathbf{F}} = [\widehat{F}_{\boldsymbol{\theta}}] \tag{7}$$

$$\mathbf{f} = [f_{\mathbf{x}}] \tag{8}$$

$$\widehat{\mathbf{f}} = [\widehat{f}_{\mathbf{x}}]. \tag{9}$$

In addition, we would also need to discretize our model for $F(\boldsymbol{\theta})$ introduced before. In order to avoid the mathematical problems associated with its possible infinitesimal structure and the need for an infinite large $\mathbf{x}$ domain, we will define it in terms of the inverse DFT of $\mathbf{f}$, namely,

$$\mathbf{F} = [F_{\boldsymbol{\theta}}] = \mathbf{M}^{-1}\mathbf{f}. \tag{10}$$

Notice we have used the term "vector" to describe the sets of field samples, although they are actually two dimensional, corresponding to the two dimensions of the aperture. This allows us to keep the language simple. We will use the term "matrix" only for the linear operators and covariance matrices defined next, which are actually four-dimensional matrices. The doubling of the dimensions are shown explicitly by the use of a two-dimensional bold index. The index $\boldsymbol{\theta}$ stands for evenly distributed values of $\sin\theta_1$ and $\sin\theta_2$, and the index $\mathbf{x}$ stands for evenly distributed values of kx_1 and kx_2. The notation is also valid for one-dimensional imaging by simply reducing the indices to a single dimension. We will use bold letter symbols for vectors and bold sans serif for matrices. We will use indexed terms within brackets whenever we want to show explicitly the contents of a vector or a matrix.

In matrix notation, equations (1) - (4) introduced for the pinhole camera transformations become

$$\mathbf{f} = \mathbf{MF} \tag{11}$$

$$\widehat{\mathbf{f}} = \mathbf{af} \tag{12}$$

$$\widehat{\mathbf{F}} = \mathbf{M}^{-1}\widehat{\mathbf{f}}, \tag{13}$$

where $\mathbf{M}$ and $\mathbf{M}^{-1}$ stand for the DFT kernel matrix $[\exp(-i\boldsymbol{\theta}\cdot\mathbf{x})]$ and its inverse, $[\exp(i\boldsymbol{\theta}\cdot\mathbf{x})]$, respectively. The effect of the aperture has been modeled above by a linear operation involving a diagonal matrix $\mathbf{a} = [a_{\mathbf{x}'}\delta_{\mathbf{x}'\mathbf{x}}]$. Here $\mathbf{x}'$ is a dummy index which will collapse in the operations involving $\mathbf{a}$. We no longer need to assume that the $a_{\mathbf{x}}$ terms are either zero or one. We will allow them to take an arbitrary complex value. Their amplitude and phase represent the weight and relative phase associated to each sampling (antenna) element of the aperture array. This will become useful later when we talk about nonredundant arrays, lenses, system inaccuracies, and array steering. By substitution of equations (11) and (12) into (13), we can write the equivalent of the pinhole camera equation (6) as

$$\widehat{\mathbf{F}} = \mathbf{AF}, \tag{14}$$

where $\mathbf{A}$ is the matrix DFT of $\mathbf{a}$, i.e.,

$$\mathbf{A} = \mathbf{M}^{-1}\mathbf{aM}. \tag{15}$$

Notice that because of the diagonal nature of $\mathbf{a}$ and the properties of the DFT operator $\mathbf{M}$, the matrix $\mathbf{A}$ is Toeplitz; that is, its elements have the property $[A_{\boldsymbol{\theta}\widehat{\boldsymbol{\theta}}}] = [A_{\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}}]$. Furthermore, we can associate a single-dimensional vector to both matrices, such that one is the DFT of the other, i.e.,

$$[a_{\mathbf{x}}] = \mathbf{M}[A_{\boldsymbol{\theta}}]. \tag{16}$$

Here $[a_{\mathbf{x}}]$ and$[A_{\boldsymbol{\theta}}]$ stand for the principal diagonal and the cross-diagonal set of terms of the matrix $\mathbf{a}$ and $\mathbf{A}$, respectively. Equation (14), then, is actually a convolving operation in matrix notation, as it should be based on the discrete convolution theorem.

3. The Statistics of Imaging: Brightness and Visibility

Since in our application, and in most other imaging applications, the object electromagnetic field is a random process in both space and time, we are not interested in particular realizations of this field but rather on its statistics. We are mainly interested in a good estimate of the angular power spectrum

$$B(\boldsymbol{\theta}) = \langle F(\boldsymbol{\theta})F^*(\boldsymbol{\theta})\rangle, \tag{17}$$

also called the "brightness distribution" by radio astronomers [e.g., *van Shooneveld,* 1979a], and more specifically in a discrete representation of it: $[B_{\boldsymbol{\theta}}] = [\langle F_{\boldsymbol{\theta}}F^*_{\boldsymbol{\theta}}\rangle]$.

$[B_{\boldsymbol{\theta}}]$ as defined above represents only the diagonal of the more general statistical matrix $\langle\mathbf{FF}^\dagger\rangle$. Here, and in what follows, the cross symbol stands for the vector or matrix adjoint, i.e., for the complex conjugate of the transpose vector or matrix.

For all practical purposes, $F(\boldsymbol{\theta})$ can be considered a finite white nonhomogeneous random process (spatial analog to nonstationary in time processes) of finite angular length. Its "whiteness" comes from the smallness of its structure, and its finite length is due to the finite illumination of the transmitter antenna and the receiver antenna element pattern (in any case, $|\boldsymbol{\theta}|$ is never bigger than 1). Therefore $\langle\mathbf{FF}^\dagger\rangle$ is a diagonal matrix, and $[B_{\boldsymbol{\theta}}]$ defines it completely.

The process is non-homogeneous because its average properties given by $B_{\boldsymbol{\theta}}$ depend on their position $\boldsymbol{\theta}$.

We have to recognize at this time two length scales: (1) the length scale of the random nature of the object field given by the correlation length of the scatterers and (2) the structural size of the brightness distribution. The difference is illustrated in Figure 1a by the structures shown by the continuous and the dashed line representing the two. Consistent with our far-field assumption, when we project our object to a plane or sphere at infinity, any finite size that the scatterer correlation length may have is projected into a practically infinitesimal angle. On the other hand, we are assuming there is a finite scale size in the structure of the brightness distribution, which we are trying to image. Therefore the first length scale is much smaller than the second. It is this relative smallness that allows us to model the object field as a white process and its statistical matrix $\langle \mathbf{F}\mathbf{F}^\dagger \rangle$ as diagonal.

Notice that the smallness of the angular correlation length with respect to the brightness structure, and therefore the diagonal nature of $\langle \mathbf{F}\mathbf{F}^\dagger \rangle$, could be questioned in the case of partial reflection structures or very aspect sensitive scatterers, especially when the object is in the near field. Such cases may require a special treatment not included here. Our treatment would still be a valid approximation if the characteristic width of the nontrivial diagonal band of the $\langle \mathbf{F}\mathbf{F}^\dagger \rangle$ matrix is smaller than the angular resolution determined by the aperture.

The whiteness of the object field implies that any one point in the aperture plane receives radiation from a large (infinite in the limit) number of independent scatterers. The large distance between the object and the aperture plane implies that all points in the latter are subject to the same radiation, except for the phases of the independent contributions. This makes the radiation field at the aperture plane, $f(\mathbf{x})$, a wide sense homogeneous Gaussian process (based on the central limit theorem). As such, it is fully characterized by its spatial autocorrelation function,

$$v(\mathbf{r}) = \langle f(\mathbf{x}) f^*(\mathbf{x}') \rangle, \tag{18}$$

where $\mathbf{r} = \mathbf{x}' - \mathbf{x}$. The autocorrelation function $v(\mathbf{r})$ is called visibility function in radio astronomy jargon [e.g., *van Shooneveld,* 1979a].

In matrix representation, the full statistical property of the field at the aperture plane is given by $\langle \mathbf{f}\mathbf{f}^\dagger \rangle$. The linear relationship between the object and the aperture fields expressed by equation (11) allows us to also express $\langle \mathbf{f}\mathbf{f}^\dagger \rangle$ linearly in terms of $\langle \mathbf{F}\mathbf{F}^\dagger \rangle$, namely,

$$\langle \mathbf{f}\mathbf{f}^\dagger \rangle = \mathbf{M} \langle \mathbf{F}\mathbf{F}^\dagger \rangle \mathbf{M}^\dagger \tag{19}$$

Given the diagonal nature of

$$\langle \mathbf{F}\mathbf{F}^\dagger \rangle = [B_{\boldsymbol{\theta}} \delta_{\boldsymbol{\theta}\boldsymbol{\theta}'}] \tag{20}$$

and the exponential nature of the Fourier transform kernel, it follows that $\langle \mathbf{f}\mathbf{f}^\dagger \rangle$ is Toeplitz; that is, it can be expressed in terms which depend only on the difference of the matrix indices as

$$\langle \mathbf{f}\mathbf{f}^\dagger \rangle = [v_{\mathbf{x}-\mathbf{x}'}]. \tag{21}$$

The Toeplitz nature of $\langle \mathbf{f}\mathbf{f}^\dagger \rangle$ corroborates our earlier deduction about the homogeneity of the process represented by $\mathbf{f}$, at least in the restricted sense.

The matrix subset $[B_{\boldsymbol{\theta}}]$ and $[v_{\mathbf{r}}]$ have dimensions N, i.e., the dimension of a "vector" (the quotes remind us of the possible two dimensionality of the $\mathbf{r}$ and $\boldsymbol{\theta}$ indices), and we will represent them by $\mathbf{B}$ and $\mathbf{v}$, respectively. They correspond to the discrete representation of the brightness and visibility function. It follows from (19) that they also form a DFT pair, i.e.,

$$\mathbf{v} = \mathbf{M}\mathbf{B}, \tag{22}$$

if we use $\mathbf{r}$ and $\boldsymbol{\theta}$ for the indices of the DFT matrix, $\mathbf{M}$. Thus not only the field functions are DFT of each other but their characterizing statistical functions are as well. This is not surprising since they correspond to an (spatial) autocorrelation function and a (angular) power spectrum. It is just the well-known Wiener-Khinchin theorem which follows from equation (19).

We can also write directly the relationship of the statistics of the field immediately below the aperture (antenna) plane in terms of the one above. It is given by

$$\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^\dagger \rangle = \mathbf{a} \langle \mathbf{f}\mathbf{f}^\dagger \rangle \mathbf{a}^\dagger, \tag{23}$$

or, combining (19) and (23), by

$$\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^\dagger \rangle = \mathbf{a}\mathbf{M} \langle \mathbf{F}\mathbf{F}^\dagger \rangle \mathbf{M}^\dagger \mathbf{a}^\dagger. \tag{24}$$

In the usual case that the diagonal matrix $\mathbf{a}$ has either 0 or 1 as elements, the matrix $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^\dagger \rangle$ is identical to the $\langle \mathbf{f}\mathbf{f}^\dagger \rangle$ matrix in a submatrix masked by the nonzero elements of $[a_{\mathbf{x}} a_{\mathbf{x}'}{}^*]$ and zero elsewhere.

In general, including cases where $[a_{\mathbf{x}}]$ are arbitrary complex coefficients, which we will consider later, the elements of $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^{\dagger} \rangle$ are given by

$$[\langle \hat{f}_{\mathbf{x}} \hat{f}^{*}_{\mathbf{x}'} \rangle] = [\langle f_{\mathbf{x}} f^{*}_{\mathbf{x}'} \rangle a_{\mathbf{x}} a^{*}_{\mathbf{x}'}] \tag{25}$$

in the nonzero submatrix at the center. It is clear, then, that in either case the $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^{\dagger} \rangle$ matrix is no longer Toeplitz and its matrix DFT

$$\langle \hat{\mathbf{F}}\hat{\mathbf{F}}^{\dagger} \rangle = \mathbf{M}^{-1} \langle \hat{\mathbf{f}}\hat{\mathbf{f}}^{\dagger} \rangle \mathbf{M}, \tag{26}$$

representing the statistics at the camera image plane, is no longer diagonal. The latter can also be expressed in terms of the object field statistics,

$$\langle \hat{\mathbf{F}}\hat{\mathbf{F}}^{\dagger} \rangle = \mathbf{A} \langle \mathbf{F}\mathbf{F}^{\dagger} \rangle \mathbf{A}^{\dagger}, \tag{27}$$

with elements

$$[\langle \hat{F}_{\hat{\boldsymbol{\theta}}} \hat{F}^{*}_{\hat{\boldsymbol{\theta}}'} \rangle] = [A_{\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}}][\langle F_{\boldsymbol{\theta}} F^{*}_{\boldsymbol{\theta}'} \rangle][A^{*}_{\boldsymbol{\theta}'-\hat{\boldsymbol{\theta}}'}] \tag{28}$$

Nevertheless, its diagonal elements,

$$\hat{B}_{\hat{\boldsymbol{\theta}}} = \langle \hat{F}_{\hat{\boldsymbol{\theta}}} \hat{F}^{*}_{\hat{\boldsymbol{\theta}}} \rangle, \tag{29}$$

continue to play the most important role. It is the equivalent to a camera picture, i.e., the equivalent to the energy deposited in the sensitive film at the focal plane of an optical (pinhole) camera. As in $\mathbf{B}$, we can define an N-dimensional vector $\hat{\mathbf{B}}$. Its relationship with the actual brightness distribution, $\mathbf{B}$, is

$$\hat{\mathbf{B}} = \mathbf{A^2}\,\mathbf{B} \tag{30}$$

where $\mathbf{A^2}$ is a Toeplitz matrix operator with terms given by

$$\mathbf{A^2} = [A_{\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}} A^{*}_{\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}}]. \tag{31}$$

Therefore equation (30) represents a convolving operation. In the particular case that the aperture is sampled by an evenly spaced square array of identical antenna elements, with a square aperture, $\mathbf{a}$ is real and symmetric, and therefore $A_{\boldsymbol{\theta}}$ is also real and symmetric. Furthermore, the elements of $\mathbf{A^2}$ are the square of the elements of $\mathbf{A}$, ${A_{\boldsymbol{\theta}}}^2$. This justifies the superscript 2 in the symbol for $\mathbf{A^2}$.

For non-Toeplitz matrices, there is an alternative definition (to the one used before) for the autocorrelation function of a process, which we can use for the truncated field $\hat{\mathbf{f}}$. We will define a cross-correlation vector $\hat{\mathbf{v}}$ in terms of the inner product of $\hat{\mathbf{f}} = [\hat{f}_{\mathbf{x}}]$ with a displaced reproduction of itself $\hat{\mathbf{f}}_{\mathbf{r}} = [\hat{f}_{\mathbf{x}+\mathbf{r}}]$, namely,

$$\hat{\mathbf{v}} = [\hat{v}_{\mathbf{r}}] = [\langle \hat{\mathbf{f}} \cdot \hat{\mathbf{f}}^{*}_{\mathbf{r}} \rangle / N]. \tag{32}$$

Notice that the index $\mathbf{x}+\mathbf{r}$ can be larger than the domain bounded by n and m. Thus, for the definition to be useful, we have to periodize $\hat{\mathbf{f}}$ with periods n and m in a checkerboard fashion. Then, it can be shown that $\hat{\mathbf{v}}$ and $\hat{\mathbf{B}}$ also form a transform pair, i.e.,

$$\hat{\mathbf{v}} = \mathbf{M}\hat{\mathbf{B}}. \tag{33}$$

However, this periodizing could also bring us some problems, especially when we will later try to reduce the size of the domain of $\mathbf{x}$ in order to reduce the number of computations. Because of the periodicity, the definition of $\hat{\mathbf{v}}$ in (32) folds values corresponding to a displacement $(\alpha m, \beta n) - \mathbf{r}$ into the desired displacement $\mathbf{r}$, where α and β can take any integer value. This aliasing effect is harmless if we keep the size of the domain at least equal to twice the size of the aperture, since the folded values would be zero. More precisely, m and n should be larger than or equal to $(2m_a - 1)$ and $(2n_a - 1)$, respectively, where m_a and n_a are the maximum projected dimensions of the aperture. This condition could be relaxed if the sizes of the intrinsic nontrivial values of the visibility function themselves are smaller than the aperture.

Let N_a be the number of nonzero elements $\hat{\mathbf{f}}_a$ of $\hat{\mathbf{f}}$. Since $\hat{\mathbf{f}}_a$ and $\mathbf{f}$ are linearly related, $\hat{\mathbf{f}}_a$ is also a N_a-variate Gaussian process. Its Gaussian nature implies that we can write its full N_a- dimensional probability distribution function in terms of N_a^2 nontrivial covariance matrix terms implicit in $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^{\dagger} \rangle$. This allows us to evaluate higher-order statistical properties of the process $\hat{\mathbf{f}}$ in terms of members of the same second-order matrix. This is important in the evaluation of errors in any brightness distribution estimate derived from $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^{\dagger} \rangle$, as well as in determining optimum estimation algorithms. Both require fourth-order expectations, but these can be expressed in terms of second- order expectations. Note that the same can be said about the full N-variate process $\hat{\mathbf{f}}$, including its zero members; we just have to include the trivial zero values of the $N \times N$ covariance matrix $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^{\dagger} \rangle$ in the discussion.

It should be recalled that any linear operation on a Gaussian process generates another Gaussian process (although not necessarily homogeneous). Therefore $\hat{\mathbf{F}}$ and any other linearly derived process from $\hat{\mathbf{f}}$ is also Gaussian, and their higher-order statistics can

be expressed in terms of the members of their own second-order statistical Matrix, like $\langle \hat{\mathbf{F}}\hat{\mathbf{F}}^{\dagger} \rangle$. This justifies including and limiting our discussion to second-order statistics.

4. Imaging Estimation Algorithms

We can express our imaging radar estimation problem as follows: Given a sampled representation of the backscattered radiation field in a finite area of the aperture plane, $\hat{\mathbf{f}}$ (i.e., the output of a finite number of antennas), estimate the brightness distribution $B(\boldsymbol{\theta})$ corresponding to a particular radar range.

Our first step is to estimate its discrete representation, $\mathbf{B}_{\boldsymbol{\theta}}$, from which one can obtain $B(\boldsymbol{\theta})$ by proper interpolation (since we have sampled $\mathbf{f}$ at the Nyquist frequency). There are several alternatives to solve the problem. We will group them into three possible approaches: the radio camera analog, a linear inversion, and a parameter estimation approach.

4.1. The Radio Camera Analog

The optical camera analog provides us with our first algorithm: We perform an inverse discrete Fourier transformation of the sampled aperture field, $\hat{\mathbf{f}}$, and evaluate the power of each element of its inverse transform, given by $\hat{\mathbf{B}}$. The approach is represented schematically in Figure 1b. The DFT can be performed either digitally or in an analog fashion. For the analog approach, we use a Butler matrix. We are including a description of such a matrix in the appendix.

We expect that $\hat{\mathbf{B}}$ gives us a good sampled estimate of $B(\boldsymbol{\theta})$. Although the corresponding continuous algorithm works well in the optical case, it results in a biased estimator in most radio cases. The difference is due to the much larger apertures used in optics, when measured in wavelengths. The technique works well only in those cases where the brightness distribution structure is much larger than the width of the convolving function $A(\boldsymbol{\theta})^2$ or, equivalently, when the characteristic size of the visibility function is much smaller than the aperture size.

One practical question arises: How large should the number of samples, $N = m \times n$, be? That is, how many zeroes should we include in $\hat{\mathbf{f}}$? On the one hand, for our mathematical formalism to be more valid, we should make N large. On the other hand, intuition tell us that a large N adds too many zeroes into our transformation box, which provide no further information. Certainly, we do not want to burden ourselves with computations that bring us no benefit. In fact, from DFT properties we know that increasing the $\mathbf{x}$ index domain with further zeroes increases the sampling resolution of $\hat{B}(\boldsymbol{\theta})$. The structure of $\hat{B}(\boldsymbol{\theta})$, in turn, is limited by the width of the convolving function $A(\boldsymbol{\theta})^2$. Higher sampling resolution is therefore redundant. Thus the domain of the index $\mathbf{x}$ does not need to be larger than twice the size of the aperture, i.e., the size of the nontrivial part of the Fourier transform of $A(\boldsymbol{\theta})^2$. It should not be less, unless the intrinsic size of the structure of $\hat{B}(\hat{\boldsymbol{\theta}})$, determined by the structure of $B(\boldsymbol{\theta})$, is wider than $A(\boldsymbol{\theta})^2$. Taking less could effectively undersample $\hat{B}(\hat{\boldsymbol{\theta}})$, losing available information. These are the same limits we discussed in the previous section to avoid aliasing in $\hat{\mathbf{v}}$.

Even if we take the proper number of terms in the transformation, we can qualify the above algorithm as a naive estimator. Because $\mathbf{f}$ is the DFT of $\mathbf{F}$, one is naively ignoring the effect of the aperture and is inverting $\hat{\mathbf{f}}$ by applying an inverse DFT, hoping it will provide a good estimator of $\mathbf{B}$.

4.2. The Linear Inversion Approach

One could expect, taking advantage of the linear relationship between $\langle \hat{\mathbf{f}}\hat{\mathbf{f}}^{\dagger} \rangle$ and $\langle \mathbf{F}\mathbf{F}^{\dagger} \rangle$ given by equation (24), to be able to obtain $\langle \mathbf{F}\mathbf{F}^{\dagger} \rangle$ by direct linear inversion. However, we are faced with a problem: The matrix $\mathbf{a}$ is singular, because of its many zero columns and rows, and cannot be inverted. Nevertheless, it is possible to do the inversion in certain cases, in an approximate way.

Let us assume that the region where

$$\langle \mathbf{f}\,\mathbf{f}^{\dagger} \rangle = [v_{\mathbf{x}-\mathbf{x}'}] \approx 0$$

can be made exactly equal to zero without any practical effect in the corresponding brightness distribution. Having done so, the nonzero region forms a diagonal band centered on the diagonal of the $\langle \mathbf{f}\,\mathbf{f}^{\dagger} \rangle$ matrix. Given its Toeplitz nature, all the terms on diagonal lines parallel to the center diagonal have the same index difference $\mathbf{r} = \mathbf{x} - \mathbf{x}'$, and therefore they take the same value given by $\langle f_{\mathbf{x}} f^{*}_{\mathbf{x}'} \rangle = v_{\mathbf{x}-\mathbf{x}'}$. Thus we need to evaluate only one element in any of these diagonal lines to obtain the corresponding value of $v_{\mathbf{x}-\mathbf{x}'}$ and of all other terms of the matrix in the same line, no matter how large is N. It is obvious, of course, that if more than one term is evaluated,

we can average them to get a statistically better estimate of the same value.

Within the matrix $\langle \mathbf{f}\,\mathbf{f}^{\dagger} \rangle$, we can identify a central "square area" (actually a submatrix of four dimensions) corresponding to the same positions where all the values of the elements of $\mathbf{a}$ are nonzero. Within this submatrix, equation (25) holds, and all of its terms can be evaluated from the terms of $\langle \hat{f}_{\mathbf{x}} \hat{f}^{*}_{\mathbf{x}'} \rangle$ by

$$\langle f_{\mathbf{x}} f^{*}_{\mathbf{x}'} \rangle = \langle \hat{f}_{\mathbf{x}} \hat{f}^{*}_{\mathbf{x}'} \rangle / (a_{\mathbf{x}} a_{\mathbf{x}'}). \tag{34}$$

Let $\mathbf{r}_{\max}$ be the largest index $\mathbf{r}$ can take with a nonzero value of $v_{\mathbf{r}}$. Now we should consider two cases: Either the submatrix is larger than the nonzero band or it is smaller. If larger, we say that $[v_{\mathbf{r}}]$ has not been truncated by the aperture. We would have obtained at least one estimate of $v_{\mathbf{r}_{\max}}$ and proportionally more estimates for the values of $v_{\mathbf{r}}$ with smaller values of $\mathbf{r}$. Our problem would be solved. We would have all nontrivial values of $[v_{\mathbf{r}}]$, and $[B_{\boldsymbol{\theta}}]$ could be obtained by transformation using equation (22). Having $\mathbf{v}$, we could also construct $\langle \mathbf{f}\,\mathbf{f}^{\dagger} \rangle$ for the whole domain of indices $\mathbf{x}\mathbf{x}'$.

If the submatrix is smaller than the band of nontrivial values of $[v_{\mathbf{r}}]$, we say that the visibility function has been truncated by the aperture and we have an incomplete solution. It is still possible to find a solution to our problem, provided we make further assumptions or have additional information. This is the subject of the next subsection.

In either of the above two cases, we need to evaluate experimentally the values of the $\langle \hat{\mathbf{f}}\,\hat{\mathbf{f}}^{\dagger} \rangle$ matrix terms, $\langle \hat{f}_{\mathbf{x}} \hat{f}^{*}_{\mathbf{x}'} \rangle$. In fact, this evaluation is the one that takes the most effort, since it involves the averaging, preferably in real time, of tens to hundreds of thousands of these products per second. This is to be compared with the inversion operations that are made once per integration period, i.e., once in times of the order of one to a few minutes. We can estimate the terms of $\langle \hat{\mathbf{f}}\,\hat{\mathbf{f}}^{\dagger} \rangle$ using two alternatives: (1) by directly evaluating each nonzero element of the matrix $\langle \hat{\mathbf{f}}\,\hat{\mathbf{f}}^{\dagger} \rangle$, $\langle \hat{f}_{\mathbf{x}} \hat{f}^{*}_{\mathbf{x}'} \rangle$, i.e., by cross correlating the corresponding outputs of the sampling antenna elements of the array, or (2) through DFTs using fast Fourier transform algorithms. The second alternative involves several steps: (1) Apply a inverse DFT, using a FFT routine, to each $\hat{\mathbf{f}}$ to obtain $\hat{\mathbf{F}}$;(2) estimate $\hat{\mathbf{B}}$ by squaring and averaging several realizations of them;(3) from $\hat{\mathbf{B}}$, obtain $\hat{\mathbf{v}}$ via equation (33); and (4) solve equation (32) to obtain the elements $\langle \hat{f}_{\mathbf{x}} \hat{f}^{*}_{\mathbf{x}'} \rangle$ of $\langle \hat{\mathbf{f}}\,\hat{\mathbf{f}}^{\dagger} \rangle$.

We should note here that equation (32) could be linearly inverted to obtain the elements $\langle \hat{f}_{\mathbf{x}} \hat{f}^{*}_{\mathbf{x}'} \rangle$ of $\langle \hat{\mathbf{f}}\,\hat{\mathbf{f}}^{\dagger} \rangle$, only if the conditions on the size N of the domain, discussed in the previous section, are satisfied. Only then, would we have $(2m_a - 1)(2n_a - 1)$ equations necessary to obtain the equal amount of independent (real and imaginary) terms in the matrix $\langle \hat{\mathbf{f}}\,\hat{\mathbf{f}}^{\dagger} \rangle$. Otherwise, the folded values could not be unfolded. This is in agreement with our previous evaluation of the amount of zero padding needed in order not to undersample $\hat{B}(\hat{\boldsymbol{\theta}})$. The importance of zero padding in preserving the amount of information is often overlooked in related spectral evaluations.

The only reason we would use the second alternative is the hope that going through a FFT would require fewer mathematical operations. We should recall that it takes only $N \log N$ operations to perform a FFT compared with $N(N-1)$ to evaluate a cross-correlation matrix. Furthermore, there is the possibility to evaluate the DFT instantaneously in an analog way by means of a Butler matrix. The last two operations, 3 and 4 above, are performed infrequently, only once per imaging experiment.

One could be tempted to stop at step 2 above and use $\hat{\mathbf{B}}$ as an estimator of $\mathbf{B}$. If we do, we would revert to our first naive estimator. In contrast, if we transform back, as suggested, we are allowed to use (34) to remove the effects of $\mathbf{a}$. This is effectively a deconvolving operation. It deconvolves the effects of $\mathbf{A}^2$.

4.3. Parameter Estimation Approach: Modeling and Maximum Entropy Method

Often we have the case of a truncated visibility function mentioned above. In this case, we could still estimate $\langle \hat{\mathbf{f}}\,\hat{\mathbf{f}}^{\dagger} \rangle$ using the linear inversion approach described in the previous subsection, but only for values of $\mathbf{r}$ for which there is at least one nonzero value of a corresponding member $a_{\mathbf{x}} a^{*}_{\mathbf{x}'}$. However, a truncated estimate of the visibility function does not allow a direct inversion to obtain the desired brightness distribution. Nevertheless, we have at our disposal many inversion techniques based on modeling the brightness distribution that conform to our (incomplete) estimates of the visibility function, provided we can apply other constraints. Possible constraints include any prior knowledge we may have about the brightness distribution and its positive-real proper-

ties and that the entropy of the model be a maximum (least number of assumptions beyond what we actually know). The latter leads us to the very popular maximum entropy techniques. It is not our intention to review them here. Many good articles and books [e.g., *Smith and Grandy,* 1985; *Skilling and Sibisi,* 1996] have been written on the subject. Many are motivated by imaging in radio astronomy and, before then, by time series frequency spectral estimation. Further discussions applied to coherent radar imaging is given by *Hysell* [1996].

Effectively, the task is to extrapolate a tail to the known part of the visibility function in a way that conforms with the positive definite properties of an autocorrelation function and with whatever a priori information we have about the object field. This is to be contrasted with the window weighting algorithms (Hanning windows and others) of *Blackman and Tukey* [1959]. Window weighting introduces further biases in our estimates which narrow the correlation function even further, with the consequential further widening of any sharp object that we might be trying to image. Even a hand-drawn extrapolation of the tail would do better than any of the different weighting windows that have been proposed.

The possibilities of parameter estimation through modeling can be illustrated with a simple example. Let us assume that we have a priori knowledge that there is a single scattering blob in the observable area in the object plane that we would like to image. That is, let us assume the brightness distribution is bell-shaped and that a Gaussian is sufficiently close to model it for our purposes. We then model the brightness distribution analytically, leaving free the six parameters that define a two-dimensional Gaussian. These parameters correspond to the blob scattering strength, its two dimensional position, its width along the major and minor axis of its elliptical cross section, and the orientation of these axis. A given set of parameters uniquely determine the shape of the visibility function for any displacement $\mathbf{r}$. On the observational side, let us assume we have only four antennas in a 2×2 square array. With them we can estimate a 4×4 element cross correlation and power matrix, $\langle \widehat{\mathbf{f}}\widehat{\mathbf{f}}^{\dagger} \rangle$, but with only 9 $(= (2 \times 2 - 1)(2 \times 2 - 1))$ degrees of freedom. We can use a least squares fitting algorithm to minimize the difference between the predicted shape of the visibility function and the few measured points. No matter how narrow the blob is, i.e., no matter how wide the visibility function is as compared with the maximum $\mathbf{r}$ we can measure, we should be able to obtain the six parameters that define the brightness distribution. Therefore the resolution of the instrument, using this technique, is not determined by the size of the array as in the radio camera approach. It is also clear that a larger number of elements would allow us to fit a larger number of independent sources. On the other hand, a linear interferometer [e.g., *Farley and Ierkic,* 1981] with only two antennas is the simplest array we can have. It can be used to determine up to three independent parameters of the model we decide to use. Recently, *Huang et al.* [1995] has determined five simple two-dimensional model parameters of middle latitude ionospheric E region irregularities using three antennas to form a two-dimensional interferometer.

Although the basic philosophy of the maximum entropy method is to minimize the number of assumptions needed to find a brightness distribution consistent with our data and with the maximum definite positiveness property of the brightness-visibility Fourier transform pair, it can also be interpreted as a fitting procedure of functions of known shape. Maximizing an entropy of the form $1/(2\theta) \int_{-\theta}^{\theta} \log B(\theta) d\theta$ can be interpreted as a fitting of $N - 1$ Lorentzian shapes, of different widths and location, to $B(\theta)$ [e.g., *van Schooneveld,* 1979b].

5. Imaging the Near Field: Focusing

The representation of the field at the aperture plane as a Fourier integral of the field at the object plane can be interpreted as a superposition of plane waves with phase given by the kernel $\exp(-\imath\boldsymbol{\theta} \cdot \mathbf{x})$ weighted by their amplitude $F(\boldsymbol{\theta})$. The amplitude $F(\boldsymbol{\theta})$, in turn, is proportional to the field at the point in the source specified by the angular coordinate $\boldsymbol{\theta}$. There is an implicit assumption in the form of the kernel that the waves are planar. This is justified by our assumptions about the object plane being in the far field. If we relax this assumption but assume that the distance is still large enough for the curvature of the field to be expressed in terms of a quadratic surface within the aperture area, we can analogously represent the field at the aperture plane by the following integral:

$$f(\mathbf{x}) = c \int \exp(-\imath\boldsymbol{\theta}\cdot\mathbf{x}) \exp(-\imath\mathbf{x}\cdot\mathbf{x}/R) F(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (35)$$

Here R is measured in wave radians and represents the distance to the object plane (actually a curved surface) specified by the radar range. The second exponential is the phase correction due to the curvature of the field at the aperture. It is a function of only $\mathbf{x}$ and can be taken outside the integral. It plays a similar role in the mathematical model as $a(\mathbf{x})$. Let us define $b(\mathbf{x}) = \exp(-\imath\mathbf{x}\cdot\mathbf{x}/R)$. We can experimentally modify our original aperture function, $a(\mathbf{x})$, designed for objects in the far field by multiplying it by $b(\mathbf{x})^{-1}$ and create a new aperture function given by $a(\mathbf{x})b(\mathbf{x})^{-1}$. With this new aperture the field right below the aperture plane is given by

$$\widehat{f}(\mathbf{x}) = a(\mathbf{x})b(\mathbf{x})^{-1}b(\mathbf{x})f(\mathbf{x}) = a(\mathbf{x})f(\mathbf{x}) \qquad (36)$$

and we are back to the equation we had before. Therefore all our previous discussions hold for this case as well.

With the function $b(\mathbf{x})^{-1}$ we have effectively introduced a radio lens to collimate the radiation coming in from the near field back into plane waves. In other words, we have created a virtual image of the object farther out at infinity. We can, then, process this virtual image in the same way we process objects at the far field. We can say that we have "focused" in the near field. However, we should note that focusing does not correct for the smearing effect of a finite aperture. The image is still convolved by the aperture function $A(\boldsymbol{\theta})^2$.

The radio lenses can be implemented by introducing extra lengths in each one of the antenna elements of the aperture array in order to correct for the undesirable phase introduced by the $b(\mathbf{x})$ term. In matrix notation, the curvature of the field and the correcting "lens" can be represented by two diagonal matrices $\mathbf{b}$ and $\mathbf{b}^{-1}$ with elements given by $[b_{\mathbf{x}}\delta_{\mathbf{x}\mathbf{x}'}]$ and $[b_{\mathbf{x}}^{-1}\delta_{\mathbf{x}\mathbf{x}'}]$, respectively.

Alternatively, let

$$\widehat{\mathbf{f}}' = \mathbf{abf} \qquad (37)$$

be the sampled field without lens correction. The desired statistics can also be obtained by applying a poststatistics digital correction to $\langle\widehat{\mathbf{f}}'\widehat{\mathbf{f}}'^{\dagger}\rangle$, namely,

$$\langle\widehat{\mathbf{f}}\widehat{\mathbf{f}}^{\dagger}\rangle = \mathbf{b}^{-1}\langle\widehat{\mathbf{f}}'\widehat{\mathbf{f}}'^{\dagger}\rangle\mathbf{b} = \mathbf{a}\langle\mathbf{f}\,\mathbf{f}^{\dagger}\rangle\mathbf{a}^{*}. \qquad (38)$$

Here we have taken advantage of the commutability of $\mathbf{a}$ with $\mathbf{b}$, since they are both diagonal.

Being able to do a lens correction after the statistics have been evaluated has the advantage that it can be done digitally without computational burden, since the correction is performed only once per integration. Furthermore, in the case of a radar, the correction could be different for different ranges. At VHF, for instance, MST radars would require no correction for mesospheric heights, but it would be certainly needed at tropospheric ones.

Note that we do not need to focus the radar transmitter. All we need is a source of illumination. The phase of the transmitter is randomized by the statistical independence of the scatterers.

Having introduced a correcting matrix $\mathbf{b}^{-1}$, we can use the same formalism for other applications. It may be that the gains and phases of the different amplifiers and antennas in an imaging system may differ from the ideal constant value. We then can devise techniques to calibrate what the effects of these imperfections are in terms of a matrix $\mathbf{b}$ and apply a correction $\mathbf{b}^{-1}$ to the statistics, in the same way as in the focusing scheme.

In addition, we could purposely introduce a phase correction, linear with respect to $\mathbf{x}$ by means of a proper $\mathbf{b}^{-1}$. We would then have introduced a "prism lens" and steered the array to a new position. If done after statistics, we would be extending to multiple antenna arrays the poststatistical steering technique introduced by *Kudeki and Woodman* [1990].

6. Economizing Antennas and Signal Processing

The first possibility we have to reduce the number of antennas and computations is by reducing the sampling arrays from two to one dimension. A two-dimensional aperture is a necessary requirement for imaging horizontally isotropic (or nearly so) irregularities such as the ones studied with MST radars and wind profilers. However, ionospheric irregularities are highly anisotropic and can be represented by a one-dimensional image obtained by a one-dimensional array of antennas. One-dimensional images and arrays imply enormous saving in computational requirements. Nevertheless, it should be mentioned that fortunately, the two-dimensional applications happen to have a much slower information rate (longer correlation times). Tropospheric processes, for instance, are more than 2 orders of magnitude slower than the ionospheric E region counterpart.

We have implicitly assumed, in what has preceded, that our observational array of antennas is filling the

aperture in an evenly distributed form. When we introduced the function $a(\mathbf{x})$ in equation (3), we assumed it took a value of one throughout the aperture and zero elsewhere. Later on, when we introduced the matrix operator $\mathbf{a}$ in equation (12), we relaxed these restrictions and made it more arbitrary, but still confined the nonzero area covered by $a(\mathbf{x})$ to a relatively small area at the center of the $\mathbf{x}$ domain. We would like to take advantage of this arbitrariness and still be able to use the same mathematical formalism for more economical apertures.

Next, we will consider sampling antenna systems that are not evenly distributed or, equivalently, have several of the antennas removed from an originally evenly distributed set. The motivation for this is to economize with the reduction of the number of antennas and receivers needed to do the imaging and to reduce the number of computations required to estimate the $\langle \widehat{\mathbf{f}}\widehat{\mathbf{f}}^{\dagger} \rangle$ matrix. The system takes advantage of the Toeplitz properties of $\langle \mathbf{f}\mathbf{f}^{\dagger} \rangle$. We have explained how this is done in the previous section, when we inverted equation (23) (at least within the submatrix delimited by the nonzero values of $\mathbf{a}$). Because of the Toeplitz nature of $\langle \mathbf{f}\mathbf{f}^{\dagger} \rangle$, it is sufficient to estimate only one cross-correlation term $f_{\mathbf{x}\mathbf{x}'}$ for each possible $\mathbf{r} = \mathbf{x}' - \mathbf{x}$, to estimate $v(\mathbf{r})$, and from it the whole matrix.

The one-dimensional nonredundant spacing antenna arrays are particularly economic. These are the antenna systems that have been used at Jicamarca for the first coherent radar imaging of E and F region irregularities [*Kudeki and Sürücü*, 1991; *Hysell*, 1996; *Hysell and Woodman*, this issue]. They are designed with the goal of reducing the number of antennas with a maximum number of possible $\mathbf{r}'$. The "nonredundant" term in the name of the technique implies that there is only one pair of antennas for each spacing, $\mathbf{x} - \mathbf{x}'$. This can be achieved, but at the cost of having no pairs for some spacings. The dilemma increases with the number of elements. However, if some degree of redundancy is allowed, we can obtain all spacings and still have an economical system. Antenna arrays that approximate a nonredundant spacing can also be designed for two dimensions.

We are depicting in Figure 2, for illustrative purposes, the center submatrix of a hypothetical cross-correlation matrix $\langle \widehat{\mathbf{f}}\widehat{\mathbf{f}}^{\dagger} \rangle$ of a linear array with a function $a_x = [0, 0, ..., 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, ..., 0, 0]$. The darkest elements of the matrix correspond to the only indices where a crosscorrelation term $\langle f_x f^*_{x'} \rangle$ can be estimated. The nonredundancy has been achieved, but at the cost of missing the spacing corresponding to $r = 5$. The figure also depicts an assumed $v(r)$. It shows some degree of truncation; that is, some nontrivial values of $v(r)$ have not been sampled at larger r's. The only way we can invert a $\langle \widehat{\mathbf{f}}\widehat{\mathbf{f}}^{\dagger} \rangle$ matrix of this type, with missing $\mathbf{r}$ at both the tail and main body of $v(\mathbf{r})$, is by using the modeling techniques discussed in the previous section, including maximum entropy methods [*Kudeki and Sürücü, 1991*; *Hysell*, 1996; *Hysell and Woodman*, this issue].

From the above point of view, a filled array produces a high degree of redundancy. On first sight, redundancy appears to be superfluous. However, we are estimating the characterizing statistical parameters, and redundancy helps in reducing the errors of the estimates in a given observational time. How much redundancy is worth taking advantage of is a matter of economics. In this regard, we should consider that in many radars, we already have the antennas. The receivers can be built at a fraction of the cost of the antennas, and computer power is becoming increasingly economical, with sufficient speed to process a large number of channels in parallel.

7. Frequency Power-Spectrum Imaging

The electromagnetic radio field at the aperture plane is a random process in space and time. It results from the superposition of random time-varying signals coming from different regions of the sky. In the previous section, we have reviewed the techniques available to us to estimate only the power level coming from each angular region of the sky (brightness distribution). However, we are not limited to just this. We have taken this approach with the purpose of keeping the concepts and mathematical notation simple and with the idea of later extending the discussions and conclusions to the space-time domain.

As in the case of any random process in time, we can consider the signal coming from a particular region of the sky as the superposition of signals of different frequencies, each with its own power. We can then extend the concept of brightness and introduce $B(\boldsymbol{\theta}, \omega)$, which can be interpreted as a frequency-dependent brightness distribution. In practice, as we did in space, we need to limit and discretize the frequency domain and associate different brightness dis-

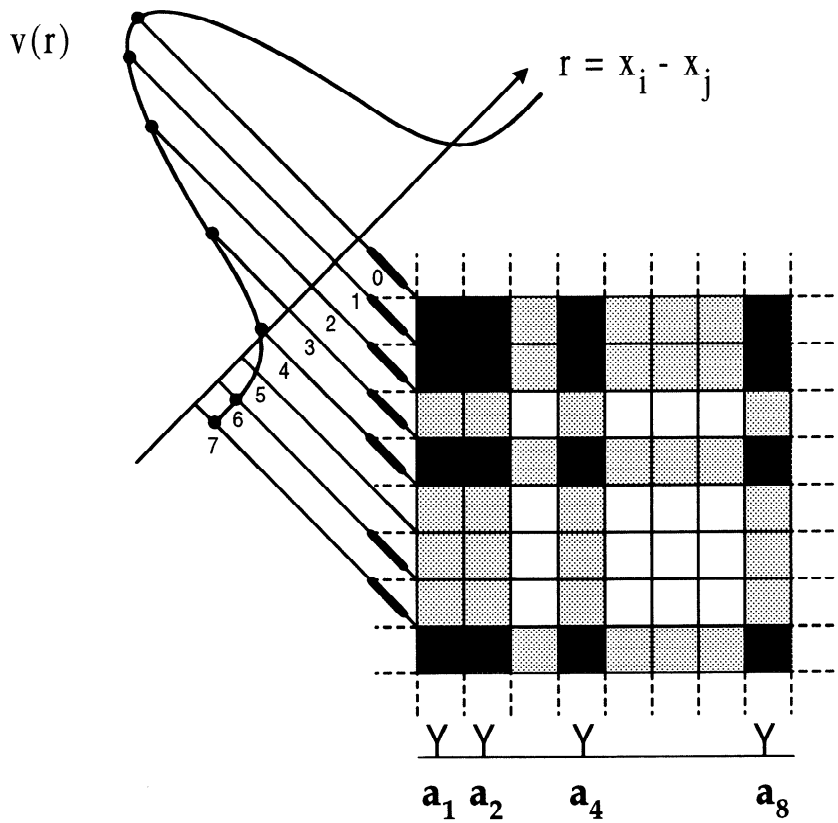

Figure 2. Graphical representation of the center part of the **a** matrix for a nonredundant spaced array as a mask to the $\langle \mathbf{f}\,\mathbf{f}^\dagger \rangle$ matrix.This could also be interpreted as the center part of the resultant $\langle \widehat{\mathbf{f}}\,\widehat{\mathbf{f}}^\dagger \rangle$ matrix. The index $\mathbf{x}$ in this case is one dimensional, corresponding to a linear array of antennas ($[a_x] = [0, ..., 0, 1, 1, 01, 0, 0, 0, 1, 0, ..., 0]$. Nonzero values of a_x are shaded, and those of $a_x a_{x'}$ are darkened. It also shows the visibility function that **a** samples in terms of the indices $r = x - x'$.

tributions to the different frequencies. This is analogous to color photography and vision, where different images are obtained at different colors (frequencies). To each frequency, there corresponds a given Doppler shift imposed by the dynamics of the medium.

All the mathematical operations transforming the fields at the different planes we have introduced in the previous sections are linear. Therefore the different frequencies forming the radiation field we are trying to image can be treated independently. They do not "talk" to each other. All of our mathematical formalism introduced before, as well as the discussions, are valid for any one frequency component and for all the components that contribute to the field. All we have to do, then, to extend our discussions to the space-frequency domain, is to replace the labeling indexes $\mathbf{x}$ by $\mathbf{x}, \omega$ and $\boldsymbol{\theta}$ by $\boldsymbol{\theta}, \omega$. The main statistical matrices become

$$\langle \mathbf{F}\mathbf{F}^\dagger \rangle = [\langle F_{\boldsymbol{\theta},\omega} F^*_{\boldsymbol{\theta}',\omega} \rangle] = [B_{\boldsymbol{\theta},\omega} \delta_{\boldsymbol{\theta}\boldsymbol{\theta}'}] \quad (39)$$

$$\langle \mathbf{f}\,\mathbf{f}^\dagger \rangle = [\langle f_{\mathbf{x},\omega} f^*_{\mathbf{x}',\omega} \rangle] \quad (40)$$

$$\langle \widehat{\mathbf{f}}\,\widehat{\mathbf{f}}^\dagger \rangle = [\langle \widehat{f}_{\mathbf{x},\omega} \widehat{f}^*_{\mathbf{x}',\omega} \rangle], \quad (41)$$

and all the equations derived for them before hold.

All of the above matrix terms shown in the right-hand brackets are cross spectra (or plain spectra when the indices $\boldsymbol{\theta}$ or $\mathbf{x}$ are the same). However, only the last one is to be evaluated experimentally, since we only have access to the signal represented by $\widehat{f}_{\mathbf{x},\omega}$. There is one for each radar range. The spectra and cross spectra can be evaluated using available

standard techniques. The terms for the other matrices, including the desired brightness distribution function, are obtained by the same imaging estimation algorithms described in previous sections.

Throughout this paper, we have assumed that the irregularities produce a backscattered field, which we would like to image. We have not been concerned about the relationship between this field and the temporal and spatial statistics of the medium that does the scattering. For this, we refer the reader to a paper on the subject by *Woodman* [1991].

Appendix: An Analog Implementation of a DFT of a Radio Field: The Butler Matrix

When we drew the radio analogy to a pinhole camera, we conceptually introduced a black box capable of performing a DFT of the sampled field under the aperture, $\widehat{f}_{\mathbf{x}}$, such as the one depicted in Figure 1b. Although we are mainly concerned in this review with the digital algorithms we have at our disposal, it is interesting to note that it is possible to construct such a black box out of interconnected transmission lines. One could then effectively construct a radio camera with analog parts. Even the visual image could be constructed by connecting the output lines of the box to light-emitting devices. The only difference with an optical camera would be the discrete nature of the output.

A "black box" capable of performing a two-dimensional DFT is known as a Butler matrix, after the name of its creator, Jesse Butler [*Butler and Lowe*, 1961]. An almost identical scheme (plus some additional ones) was presented independently by *Shelton and Kelleher* [1961]. Both were motivated to obtain multiple beams from a given array of subantennas. Nevertheless, the application that motivated them could not be called imaging properly, since it involved the switching of the available beams to accommodate a single transmitter and/or receiver.

A Butler matrix consists of a set of $N \log N$ "hybrid" circuits connected by means of cables of the appropriate length. The hybrids themselves can be implemented with cables. They are a four-port circuit with a convenient characteristic: One of the two (output) ports presents the sum, and the other presents the difference of the other two (input) ports. The scheme is shown in Figure A1a. Details are given by *Butler and Lowe* [1961], and *Shelton and Kelleher* [1961].

Even more interesting is the fact that the Butler matrix is exactly an analog implementation of the fast Fourier transform (FFT). More interesting still is the fact that it was published 5 years (!) before the famous publication of *Cooley and Tukey*, [1965] of the first digital implementation of the FFT algorithm. It appears that despite the difference in time, both developments were independent, since no reference is made in the later work to Butler's matrix. The exact correspondence between the two can be appreciated in Figure A1, where we display a figure from Butler's original paper [*Butler and Lowe*, 1961]and a graphical representation of a FFT taken from *Cochran et al.* [1967]. Butler's (four-port) hybrid should be identified by the bow tie set of lines joining a corresponding set of four points: two input splitting points connecting to two adding points. It corresponds to a $N = 2$ complex Fourier transform.

Although they do not mention it, both *Butler and Lowe* [1961], and *Shelton and Kelleher* [1961] must have realized they were implementing a discrete Fourier transform. However, it appears they were not aware they were implementing the equivalent of a fast Fourier transform (FFT) numerical algorithm capable of performing a Fourier transform with only $N \log N$ operations. Otherwise, they would have impacted the scientific community in the same way *Cooley and Tukey* [1965] impacted it 5 years later. Even today, very few scientists are aware of the correspondence of Butler's hardware scheme to the FFT algorithm, and even fewer are aware of the lead time in its discovery. In fact, we have not seen due credit in the literature to this discovery.

Although speed of processing is not an issue in an analog implementation, the Butler matrix allows an economical implementation of a N input FFT analog operator, with only $N \log N$ hybrid elements. The multiplication by a complex exponential, $W^n = \exp(-\imath n 2\pi/N)$, needed in the digital algorithm is implemented by a transmission line of proper length.

Imaging could be completed by attaching a detector at each output. Integration to obtain the statistical average of the field intensity coming from each direction θ_i could also be accomplished in an analog way, but only for passive imaging, or radar imaging at a single range. For the more usual radar case, range gating will force the use of simple digital integration schemes.

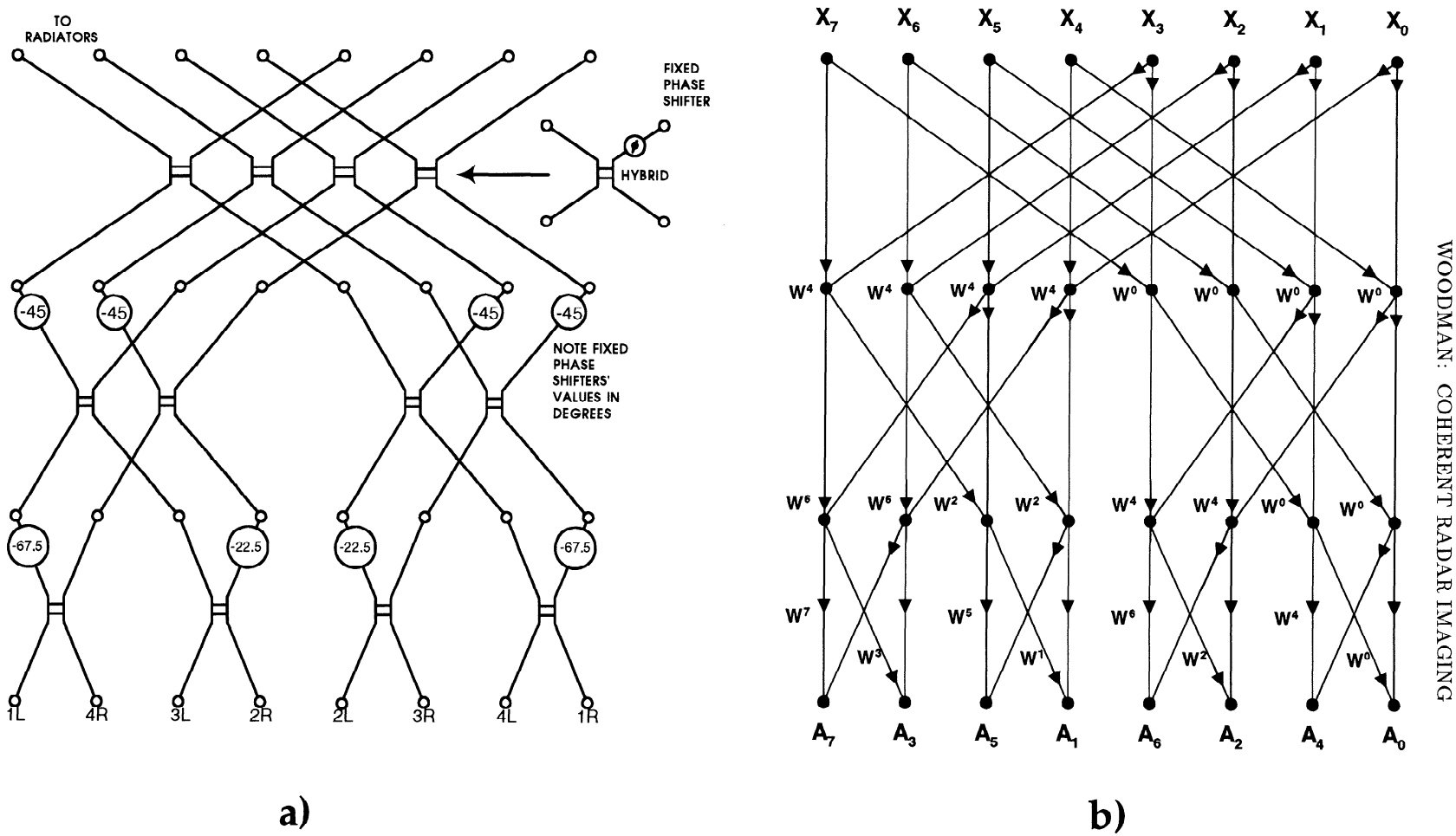

Figure A1. (a) Original diagram of a Butler matrix as presented by the author [*Butler and Lowe*, 1961] and (b) a schematic of a FFT algorithm by *Cochran et al.* [1967].

As far as we know, a Butler's matrix has yet to be used to implement radar imagining. *Greenwald et al.* [1978] has used it to obtain multiple beams in the Scandinavian Twin Auroral Radar Experiment radar, but the outputs were switched to share a single receiver. Its potential for radar imaging is there to tap. Its main advantage would be its speed of processing; being analog, it is practically instantaneous. The large number of receivers that are needed would also be required by a digital implementation of the same.

Acknowledgments. The present work has been carried out as part of a Cooperative Agreement between the Instituto Geofísico del Perú and Cornell University, partially financed by ATM-9408441. I would like to thank Jorge Chau, Don T. Farley, and the reviewers, Marku S. Lehtinen, Wes Swartz, Jürgen Röttger and an anonymous reviewer, for their comments and suggestions on the original manuscript.

References

Blackman, R.B., and J.W. Tukey, *The Measurement of Power Spectra From the Point of View of Communications Engineering*, Dover, Mineola, N.Y. 1959.

Butler, J., and R. Lowe, Beam-forming matrix simplifies design of electronically scanned antennas, *Electron. Des.*, 170-173, 1961.

Cochran, W.T, J.W. Cooley, D.L. Favin, H.D. Helms, R.A. Kaenel, W.W. Lang, G.C. Maling, D.E. Nelson, C.M. Rader, and P.D. Welch, What is the fast Fourier transform?, *IEEE Trans. Audio Electroacoust., AU-15*, 45-55, 1967.

Cooley, J. W., and J.W. Tukey, An algorithm for the machine calculation of complex Fourier series, *Math. of Comput.,19*, (90), 297-301, 1965.

Farley, D.T., H.M. Ierkic, and B.G. Fejer, Radar interferometry: A new technique for studying plasma turbulence in the ionosphere, *J. Geophys. Res., 86*, 1467-1472, 1981.

Greenwald, R.A., W. Weiss, E. Nielsen, and N.R. Thomson, STARE: A new radar auroral backscatter experiment in northern Scandinavia, *Radio Sci., 13*, 1021-1039, 1978.

Huang, C.M., E. Kudeki, S.J Franke, C.H. Liu, and J. Röttger, Brightness distribution of midlatitude *E* region echoes detected at the Chung-Li VHF radar, *J. Geophys. Res.,100*,14,703-14,715,1995.

Hysell, D.L., Radar imaging of equatorial *F* region irregularities with maximum entropy interferometry, *Radio Science*, 31, 1567-1578, 1996.

Hysell, D.L., and R.F. Woodman, Imaging coherent backscatter radar observations of topside equatorial spread *F*, *Radio Sci.*, this issue.

Kudeki E., and R.F. Woodman, A poststatistics steering technique for MST radar applications, *Radio Sci., 25*, 591-594, 1990.

Kudeki, E., and F. Sürücü, Radar interferometric imaging of field-aligned plasma irregularities in the equatorial electrojet, *Geophys. Res. Lett., 18*, 41-44, 1991.

Ostro, S.J., Planetary radar astronomy, *Rev. Mod. Phys.,65*(4) 1235-1279, 1993.

Ratcliffe, J.A., Some aspects of diffraction theory and their application to the ionosphere, *Rep. Prog. Phys., 19*, 188-267, 1956.

Shelton, J.P., and K.S. Kelleher, Multiple beams from linear arrays, *IRE Trans. Antennas Propag.*, pp. 154-161, 1961.

Skilling, J., and S. Sibisi (Eds.),*Maximum Entropy and Bayesian Methods, Proceedings of the 14th International Workshop on Maximum Entropy and Bayesian Methods, Cambridge, England, 1994*, vol. 70, *Fundamental Theory of Physics*, 323 pp., Kluwer Acad., Norwell, Mass., 1996.

Smith, C.R., and W.T. Grandy, Jr. (Eds.), *Maximum-Entropy and Bayesian Methods in Inverse Problems, Fundamental Theories of Physics*, 492 pp. Kluwer Acad., 1985.

Sürücü, 1961], F., Investigation of radar interferometric technique with applications to the atmosphere, Ph.D. thesis, Dep. of Electr. and Comput. Eng., Univ. of Ill., Urbana, 1992.

van Schooneveld, C. (Ed.), *Image Formation from Coherence Functions in Astronomy*, D. Reidel, Norwell, Mass., 1979a.

van Schooneveld, C., Resolution enhancement: The "maximum entropy method" and the "high resolution method", in *Image Frmation from Coherence Functions in Astronomy* , pp. 197-218, D. Reidel, Norwell, Mass., 1979b.

Woodman, R.F., and C. La Hoz, Radar observations of F region equatorial irregularities, *J. Geophys. Res., 81*, 5447-5466, 1976.

Woodman, R.F., A general statistical instrument theory of atmospheric and ionospheric radars, *Geophys. Res., 96*, 7911-7928, 1991.

R.F. Woodman, Instituto Geofísico del Perú, Radio Observatorio de Jicamarca, Apartado 13-0207, Lima 13, Perú. (e-mail: ron@geo.igp.gob.pe)

(Received January 28, 1997; revised June 27, 1997; accepted July 16, 1997.)