Title
Uniqueness properties of higher-order autocorrelation functions

Journal
Journal of the Optical Society of America A, 9(3)

ISSN
1084-7529

Authors
Yellott, John I
Iverson, Geoffrey J

Publication Date
1992-03-01

DOI
10.1364/josaa.9.000388

Uniqueness properties of higher-order autocorrelation functions

John I. Yellott, Jr., and Geoffrey J. Iverson

Department of Cognitive Sciences, University of California, Irvine, Irvine, California 92717

Received December 12, 1990; accepted June 12, 1991; final manuscript received October 11, 1991

The kth-order autocorrelation function of an image is formed by integrating the product of the image and k independently shifted copies of itself: The case $k = 1$ is the ordinary autocorrelation; $k = 2$ is the triple correlation. Bartelt *et al.* [Appl. Opt. **23**, 3121 (1984)] have shown that every image of finite size is uniquely determined up to translation by its triple-correlation function. We point out that this is not true in general for images of infinite size, e.g., frequency-band-limited images. Examples are given of pairs of simple band-limited periodic images and pairs of band-limited aperiodic images that are not translations of each other but that have identical triple correlations. Further examples show that for every k there are distinct band-limited images that have identical kth-order autocorrelation functions. However, certain natural subclasses of infinite images *are* uniquely determined up to translation by their triple correlations. We develop two general types of criterion for the triple correlation to have an inverse image that is unique up to translation, one based on the zeros of the image spectrum and the other based on image moments. Examples of images satisfying such criteria include diffraction-limited optical images of finite objects and finite images blurred by Gaussian point spreads.

1. INTRODUCTION

The autocorrelation of a real-valued function f is another real function a_f created by integrating the product of f and a shifted copy of itself: a function of the general form $a_f(s) = \int f(x)f(x + s)\mathrm{d}x$. It is a basic and sometimes frustrating fact of Fourier analysis that the autocorrelation function (ACF) a_f completely identifies the amplitude spectrum of f but provides no information about its phase spectrum. In recent years there has been growing interest in the possibility of recovering phase information from higher-order ACF's created by integrating the product of a function and multiple shifted copies of itself.[1-9] Generalizing the concept of autocorrelation, one can construct a sequence $\{a_{k,f}: k = 1, 2, \ldots\}$ of ACF's of a real function f, where the kth-order ACF $a_{k,f}(s_1, \ldots, s_k)$ is created by an integral of the form $\int f(x)f(x + s_1)\ldots f(x + s_k)\mathrm{d}x$. In this sequence $a_{1,f}$ is the ordinary ACF and $a_{2,f}$ is the triple-correlation function, which has been widely applied in optics[3] and is beginning to find uses in vision research.[4,10] (The Fourier transform of the triple correlation is commonly known as the bispectrum, so our numbering agrees with standard terminology in the spectral domain.)

In the optics literature, application of the triple-correlation function to phase-recovery problems is usually justified by reference to a uniqueness theorem that is due to Bartelt *et al.*[11] showing that, if a real function f has bounded support, $a_{2,f}$ determines f up to a translation [i.e., up to the form $f(x + c)$, where c is an unidentifiable centering parameter]. In other words, the triple correlation of any image of finite size contains sufficient information to identify both the amplitude spectrum and (except for a centering term) the phase spectrum of that image.

A natural question is whether this is also true for images of infinite size: frequency-band-limited images, for example, or finite images blurred by Gaussian point-spread functions. This question does not seem to have been directly addressed in the recent literature, but algorithms for recovering infinite-duration temporal signals from their triple correlations have appeared,[12] and one might be led to think that images with infinite support, like finite images, are always uniquely determined up to translation by their triple correlations.

However, simple counterexamples show that this is not the case. Figure 1 illustrates a pair of nonnegative band-limited integrable functions that have the same triple-correlation function (as is shown below in Subsection 2.D.3) but that are not translations of each other: The functions are $\mathrm{sinc}^2(x)(1 + \cos 6\pi x)$ and $\mathrm{sinc}^2(x)(1 + \sin 6\pi x)$, where $\mathrm{sinc}(x) = \sin \pi x/\pi x$. Figure 2 illustrates a pair of nonnegative periodic functions with the same property: The functions are $2 + \cos 2\pi x + \cos 6\pi x$ and $2 + \cos 2\pi x - \cos 6\pi x$. (This example is due to Klein and Tyler.[4]) More generally, it can be shown that for every k there are pairs of nonnegative band-limited integrable functions and also pairs of nonnegative band-limited periodic functions that have identical kth-order ACF's but are not translations of each other. (Examples are given in Subsections 2.D and 2.E.) Thus without the assumption of bounded support, neither the triple-correlation function nor any other finite-order ACF uniquely determines every image up to a translation: The best one can do is identify useful special classes of infinite images that are so determined.

The identification of such classes is the main purpose of this paper. We examine the uniqueness properties of higher-order ACF's of functions that represent monochromatic images: nonnegative real functions defined on the line $\mathbf{R}$ or the plane $\mathbf{R}^2$. Two general classes of image function are considered: integrable functions (i.e., $f \in L^1$) with bounded or infinite support (the latter being

0740-3232/92/030388-17$05.00

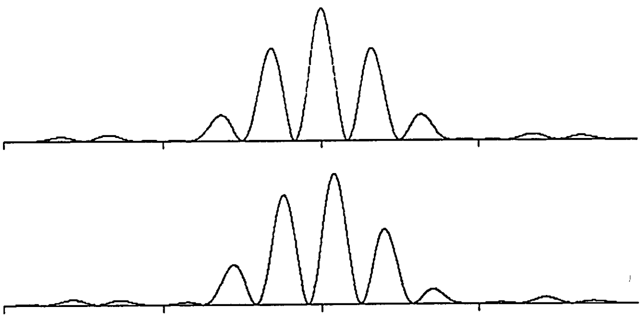

Fig. 1. Nonnegative integrable functions that have the same triple-correlation function. Top: $\text{sinc}^2(x)(1 + \cos 6\pi x)$. Bottom: $\text{sinc}^2(x)(1 + \sin 6\pi x)$.

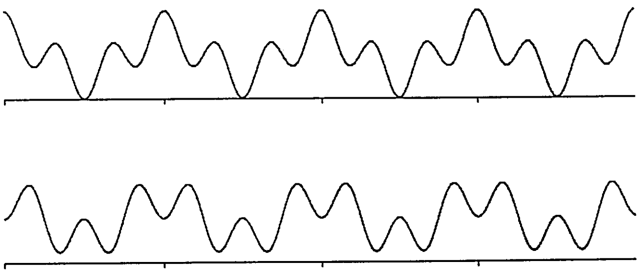

Fig. 2. Nonnegative periodic functions that have the same triple-correlation function. Top: $2 + \cos 2\pi x + \cos 6\pi x$. Bottom: $2 + \cos 2\pi x - \cos 6\pi x$.

our chief concern) and infinitely extended periodic functions. For integrable images with bounded support we re-prove the triple-correlation-uniqueness theorem of Bartelt *et al.*[11] by using a different approach, one that can be extended to certain classes of images with infinite support. The proof of Bartelt *et al.* relied on the fact that the Fourier transform of any function with bounded support is determined, up to a translation factor, by the zeros of its analytic continuation, which can be identified from the zeros of the analytic continuation of its bispectrum. That approach fails for functions with infinite support, whose complex transforms may be nonvanishing (e.g., Gaussians). Our proof is based on a functional-equation argument combined with the fact that all integrable functions with bounded support are uniquely determined by the values of their Fourier transforms in a neighborhood of the origin (that is, by the derivatives of the transform at zero, which determine the entire transform by means of a Taylor series). From a functional-equation standpoint the key fact about two images f and g with the same triple correlation is that their Fourier transforms F and G must satisfy a relationship of the form

$$F(u)F(v)F(-u - v) = G(u)G(v)G(-u - v) \tag{1}$$

for all arguments u, v. Thus the uniqueness properties of the triple correlations of images are intimately related to the solutions of Eq. (1) with F and G complex functions on $\mathbf{R}$ or $\mathbf{R}^2$. If f and g belong to a class of functions for which all solutions of Eq. (1) take the form

$$G(u) = \exp(i2\pi c \cdot u)F(u) \tag{2}$$

for some constant $c \in \mathbf{R}$ or $\mathbf{R}^2$, then f and g have the same triple correlation if and only if $g(x) = f(x + c)$. This is true for all integrable image functions with bounded support, because it can be shown that Eq. (1) implies that Eq. (2) holds for all u in some neighborhood of the origin and the transforms of finite images are determined everywhere by their values in such a neighborhood. For infinite images it remains true that Eq. (1) implies Eq. (2) in a neighborhood of the origin, but in general it is no longer the case that the image transform is completely determined by its values in such a neighborhood. Thus for infinite images one needs to impose additional constraints to guarantee that Eq. (2) is the only solution to Eq. (1).

Two kinds of constraint suggest themselves. One involves restricting the zeros of the image transform in such a way that solution (2) can be recursively extended from a neighborhood of the origin to all values of u. From an image-reconstruction standpoint, this approach corresponds to the recursive algorithms proposed by several authors for recovering the phase spectrum of an image from that of its bispectrum.[11,13–15] Such algorithms implicitly rely on the bispectrum's having adequate support, which is always guaranteed for finite images but not for infinite ones unless the zeros of the image spectrum are constrained.

The other approach is to use the fact that, when Eq. (2) holds in a neighborhood of the origin, the moments of $g(x)$ are identical to those of $f(x + c)$ for some c, which may be enough to guarantee that $g(x) = f(x + c)$. This approach to triple-correlation uniqueness corresponds to a second type of image-reconstruction algorithm, in which the moments of the image are recovered from moments of its triple correlation and the image is reconstructed from its moments. Such a reconstruction is always possible in principle for finite images but not for infinite ones unless the image moments are appropriately constrained.

Using these ideas, we show that the triple-correlation function uniquely determines, up to translation, the following types of one-dimensional (1-D) and two-dimensional (2-D) integrable image:

(1) *Images whose Fourier transforms are nonvanishing everywhere* (e.g., Gaussians, Gabor functions, exponential and gamma densities, Cauchy densities, and any convolution of such images).

(2) *Band-limited images whose transforms have no zeros* [e.g., $\text{sinc}^2(x)$, $J_1^2(r)/r^2$], *or at most a finite number of zeros* [e.g., $\text{sinc}^2(x)$ or $J_1^2(r)/r^2$ convolved with any image of finite size], *below the frequency cutoff.*[16] In two dimensions the condition refers to zeros along the axes corresponding to each spatial-frequency orientation. This result implies that the triple-correlation function determines all the diffraction-limited incoherent optical images of finite objects formed with (for example) square or circular exit pupils.

(3) *Images whose transforms have at most a finite number of zeros in every finite interval.* (In two dimensions, this refers to intervals along the axes corresponding to each spatial-frequency orientation.) Examples include any of the functions cited in condition (1) convolved with any image of finite size, in particular, any finite image blurred by any Gaussian point-spread function.

(4) *Images with the property that every point in frequency space at which the transform is not zero can be finitely linked to a neighborhood of the origin in which the transform is nonvanishing.* This is a technical condition that is not easy to state concisely but that is useful

for establishing triple-correlation uniqueness in cases in which the transform of a band-limited image vanishes over an interval below the frequency cutoff. For example, it shows that the triple correlation determines the functions $\mathrm{sinc}^2(x)(1 + \cos 5\pi x)$ and $\mathrm{sinc}^2(x)(1 + \sin 5\pi x)$, which are similar to the counterexamples of Fig. 1 except for the size of the gaps in their spectra.

(5) *Images that are uniquely determined by their moments.* The moments of an image f are $\mu_n = \int_{-\infty}^{\infty} x^n f(x)\mathrm{d}x$ in one dimension and $\mu_{n,m} = \int_{-\infty}^{\infty} x^n y^m f(x,y)\mathrm{d}y\mathrm{d}x$ in two dimensions, and f is said to be determined by its moments if there is only one nonnegative function that has the moment sequence $\{\mu_n\colon n = 0, 1, 2, \ldots\}$ or $\{\mu_{n,m}\colon n = 0, 1, 2, \ldots;\ m = 0, 1, 2, \ldots\}$. The general problem of characterizing all the functions that are determined by their moments is complicated,[17,18] but one simple sufficient condition in one dimension is $\lim_{n\to\infty} \sup|\mu_n/n!|^{1/n} < \infty$. [There is an analogous condition in two dimensions involving the quantities $M_n = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x^2 + y^2)^{n/2} f(x,y)\mathrm{d}y\mathrm{d}x, n = 0, 1, 2, \ldots$.] If the moments of an image satisfy that condition, its complex Fourier transform is analytic in a strip containing the real axis and consequently is uniquely determined for all real arguments by its derivatives at the origin, which can be obtained from certain derivatives of the bispectrum. [In that case the transform cannot have infinitely many zeros in any finite interval of the real axis, so the moments condition becomes redundant with condition (3) above, although perhaps easier to verify in some contexts.] However, there are also images that are determined by their moments and thus by their triple correlation but that do not satisfy the limit condition: $f(x) = \exp(-x^{1/2})\,(x \geq 0)$ is a 1-D example.

Thus certain natural subclasses of integrable images with infinite support *are* uniquely determined up to translation by their triple correlations. But this is not true of the entire class, and the imposition of limits on the image bandwidth does not improve the situation, because there are images with arbitrarily small bandwidths (e.g., rescalings of the images in Fig. 1) that are not determined by their triple correlations. One's next hope might be that the entire class of integrable images would be determined by the kth-order ACF for some fixed $k > 2$. But that is not the case: as was noted above, for every k one can find pairs of integrable band-limited images that are not translations of each other but whose ACF's agree for all orders 1 through k. However, one can show that, if *all* the ACF's of two integrable images f and g are identical (i.e., $a_{k,g} = a_{k,f}$ for all k), then f and g must be identical except for a translation.

We also sketch the uniqueness properties of the higher-order ACF's of infinitely extended periodic images, drawing on the work of Klein and Tyler.[4] Here again two images f and g have the same triple-correlation function if and only if their transforms satisfy Eq. (1), but the discreteness of the spectrum in this case weakens the force of that constraint, making the uniqueness problem more difficult. For example, it is easy to prove that every integrable image is determined up to translation by its entire set of ACF's of all orders, but it is not so obvious whether this is also true of periodic images. As in the integrable case, for every k one can find pairs of distinct band-limited periodic images that have the same kth-order ACF, and one can quickly show that the set of ACF's of all orders determines any periodic image whose spectrum is nonvanishing at the fundamental frequency. But we do not know whether this is true for arbitrary periodic images, and we leave that as an open problem.[19] On the positive side, we show that the triple correlation uniquely determines all periodic images that satisfy conditions analogous to condition (1) or (2) above.

The uniqueness properties of higher-order ACF's are essentially the same for 1-D and 2-D images. We deal first (in Section 2) with the 1-D case in some detail, and then (in Section 3) we show that the same reasoning can be readily extended to 2-D images. It will be seen that the proofs in Section 3 are actually independent of dimensionality, so with natural rephrasing the theorems hold for nonnegative real functions on $\mathbf{R}^n$ for any n.

From a reconstruction standpoint, the fact that some infinite images are not uniquely determined up to translation by their triple correlation raises two immediate questions, which we address in Subsection 2.F. In such cases a given triple-correlation function could have been generated by at least two disjoint families of images, say $\{f(x + c)\}$ and $\{g(x + c)\}$. What happens if one tries to invert such a triple-correlation function? Is one of the possible inverse-image families arbitrarily singled out, or does the inversion procedure simply fail to give any result? (For the two inversion procedures considered here the answer is the latter.) The other question concerns the fact that in principle any finite portion of an infinite image can be uniquely reconstructed from its triple correlation, even though the triple correlation of the entire image may not have a unique inverse. What effect does the nonuniqueness of the triple correlation of the entire image have on one's ability to reconstruct a finite portion of it? (Basically, the answer here is that reconstruction becomes increasingly unreliable as the size of the finite portion grows, because one is forced to assign definite phase values to the bispectrum at points where its absolute value is suspiciously close to zero. In a sense, the more we know about the bispectrum of such an image, the less certain we become about the image itself.)

To avoid possible misunderstanding, it is noted explicitly that, while the following analysis often relies on probabilistic arguments (exploiting the formal similarity between images and probability distributions), the images that concern us are always deterministic: the paper does not deal with higher-order ACF's of stochastic processes.

2. HIGHER-ORDER AUTOCORRELATIONS OF 1-D IMAGES

A. Definitions

We think of a function $f\colon \mathbf{R} \to \mathbf{R}$ as representing a 1-D image in the sense that the total amount of light in any interval I is $\int_I f(x)\mathrm{d}x$. We will say that f is an image function (or, simply, an image) if f is nonnegative and integrable over every finite interval.[20] If two functions f and g have the same integral for every I so that $f = g$ almost everywhere, then f and g represent physically indistinguishable images, and we write simply $f = g$. In particular, if $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 0$, then (since f is nonnegative) $f = 0$. We consider the ACF's of two classes of image: integrable and periodic. If $\int_{-\infty}^{\infty} f(x)\mathrm{d}x$ is finite, f is an integrable image and its kth-order ACF is denoted $a_{k,f}$ and defined as

$$a_{k,f}(s_1,\ldots,s_k) = \int_{-\infty}^{\infty} f(x)f(x+s_1)\ldots f(x+s_k)\mathrm{d}x. \quad (3)$$

A probabilistic argument shows that $a_{k,f}$ is integrable over $\mathbf{R}^k$ for all k. Let $\alpha = \int_{-\infty}^{\infty} f(x)\mathrm{d}x$. By definition α is finite, and it is positive unless $f = 0$, in which case $a_{k,f} = 0$ and the claim is trivially true. Assuming that $\alpha > 0$, let $p_f(x) = f(x)/\alpha$. Then p_f is a probability-density function. Let $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k$ be $k+1$ independent random variables, each having the density function p_f. Then the probability-density function of the k-dimensional random vector $(\mathbf{x}_1 - \mathbf{x}_0, \mathbf{x}_2 - \mathbf{x}_0, \ldots, \mathbf{x}_k - \mathbf{x}_0)$ at the point $(s_1,\ldots,s_k)$ is $\int_{-\infty}^{\infty} p_f(x)p_f(x+s_1)\ldots p_f(x+s_k)\mathrm{d}x$, which equals $a_{k,f}(s_1,\ldots,s_k)\alpha^{-k-1}$. Since the density of the random vector must be integrable over $\mathbf{R}^k$, so is $a_{k,f}$.

We say that f is a periodic image if $f(x) = \sum_{m=-\infty}^{\infty} f_p(x - mp)$, where $f_p(x)$ is an image function defined arbitrarily on an open interval $(-p/2, p/2)$, with $f_p(x) \equiv 0$ for $|x| \geq p/2$. The kth-order ACF of a periodic image f, denoted $\ddot{a}_{k,f}$, is defined as

$$\ddot{a}_{k,f}(s_1,\ldots,s_k) = \lim_{w\to\infty}(1/w)\int_{-w/2}^{w/2} f(x)f(x+s_1)\ldots f(x+s_k)\mathrm{d}x,$$

which is equivalent to the definition

$$\ddot{a}_{k,f}(s_1,\ldots,s_k) = (1/p) \times \int_{-p/2}^{p/2} f(x)f(x+s_1)\ldots f(x+s_k)\mathrm{d}x. \quad (4)$$

The function $\ddot{a}_{k,f}$ is periodic in $\mathbf{R}^k$. Its integral over every k-dimensional cube with sides of length p is $(1/p)\beta^{k+1}$, where β is the integral of f over a single period.

[In the first paper applying higher-order ACF's to visual perception, Klein and Tyler[4] analyzed what they called the "generalized autocorrelations" of periodic images, which are the same as the kth-order ACF's defined by Eq. (4) except for numbering: the "kth-order generalized autocorrelation" in their terminology is our $(k-1)$st-order ACF $\ddot{a}_{k-1,f}$.]

B. Fourier Transforms

The Fourier transform F of an integrable image f is defined as

$$F(u) = \int_{-\infty}^{\infty} \exp(-i2\pi ux)f(x)\mathrm{d}x.$$

The Fourier transform of its kth-order ACF $a_{k,f}$ is denoted $A_{k,f}(u_1,\ldots,u_k)$. It is related to F by the following calculation:

$$\begin{aligned} A_{k,f}(u_1,\ldots u_k) &= \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} \exp\left(-i2\pi\sum_{j=1}^{k} s_j u_j\right) \times a_{k,f}(s_1,\ldots,s_k)\mathrm{d}s_1\ldots\mathrm{d}s_k \\ &= \int_{-\infty}^{\infty} f(x)\left[\prod_{j=1}^{k}\int_{-\infty}^{\infty}\exp(-i2\pi u_j s_j) \times f(x+s_j)\mathrm{d}s_j\right]\mathrm{d}x \\ &= F\left(-\sum_{j=1}^{k} u_j\right)\prod_{j=1}^{k} F(u_i). \end{aligned} \quad (5)$$

The transforms $A_{k,f}$ are called the higher-order spectra or polyspectra of f, $A_{2,f}$ being the bispectrum. Equation (5) implies that integrable images f and g have the same triple correlation if and only if their transforms F and G satisfy Eq. (1).

If f is a periodic image with period p, its Fourier transform (in the sense of generalized function theory) is

$$F(u) = \sum_{m=-\infty}^{\infty} (1/p)F_p(m/p)\delta(u - m/p),$$

where $\delta()$ is the Dirac delta function and $F_p(u)$ is the transform of the single-period segment f_p:

$$F_p(u) = \int_{-p/2}^{p/2} \exp(-i2\pi ux)f(x)\mathrm{d}x.$$

The Fourier transform of the kth-order ACF $\ddot{a}_{k,f}$ of a periodic image f is denoted $\ddot{A}_{k,f}$. A calculation analogous to the one leading to Eq. (5) shows that

$$\ddot{A}_{k,f}(u_1,\ldots,u_n) = (1/p)F_p\left(-\sum_{j=1}^{k} u_j\right) \times \prod_{j=1}^{k}\left[\sum_{m=-\infty}^{\infty}(1/p)F_p(m/p)\delta(u_j - m/p)\right].$$

The function $\ddot{A}_{k,f}$ is concentrated at k-dimensional lattice points of the form $(u_1,\ldots,u_k) = (m_1/p,\ldots,m_k/p)$, where the m_j are integers. At those points its k-dimensional delta has the value

$$(1/p)F_p\left(-\sum_{j=1}^{k} m_j/p\right)\prod_{j=1}^{k}(1/p)F_p(m_j/p).$$

This expression is the product of the values of the deltas of $F(u)$ at $u = m_1/p,\ldots,m_k/p, -\sum_{j=1}^{k} m_j/p$, so for the periodic case the transform of the kth-order ACF of f is again determined by that of f itself through a relationship of the form of Eq. (5), i.e.,

$$\ddot{A}_{k,f}(u_1,\ldots,u_k) = F\left(-\sum_{j=1}^{k} u_j\right)\prod_{j=1}^{k} F(u_j). \quad (6)$$

Thus two periodic images f and g have the same triple correlation if and only if their transforms satisfy Eq. (1).

C. Basic Properties of Higher-Order ACF's

The following statements are true for both integrable and periodic images:

(i) The kth-order ACF determines all lower orders [i.e., if $a_{k,f} = a_{k,g}$ $(\ddot{a}_{k,f} = \ddot{a}_{k,f})$, then $a_{j,g} = a_{j,g}$ $(\ddot{a}_{j,f} = \ddot{a}_{j,g})$ for all $j < k$].

(ii) Translating an image leaves all its ACF's unchanged [i.e., if $g(x) = f(x+c)$, then $a_{k,g} = a_{k,f}$ $(\ddot{a}_{k,g} = \ddot{a}_{k,f})$ for all k].

(iii) If images f and g have the same kth-order ACF and both are convolved with any third function h, the convolutions $h * g$ and $h * f$ have the same kth-order ACF.

(iv) If images f and g have the same kth-order ACF, so do the convolutions $f * f$ and $g * g$.

(v) If images $f(x)$ and $g(x)$ have the same kth-order ACF, so do $f(ax)$ and $g(ax)$ for any constant $a \neq 0$.

Properties (i)–(v) follow immediately from Eq. (5) for integrable images or from Eq. (6) for periodic images: (i) $a_{k,f} = 0$ if $f = 0$, in which case $a_{k-1,f} = 0$. Otherwise

$F(0) > 0$, and $A_{k-1,f}$ can be obtained from $A_{k,f}$ by setting $u_k = 0$ in Eq. (5):

$$F(0)^{-1}A_{k,f}(u_1,\ldots,u_{k-1},0) = A_{k-1,f}(u_1,\ldots,u_{k-1}).$$

For $\ddot{a}_{k,f}$ the same argument holds when Eq. (6) is used. (ii) If $g(x) = f(x + c)$, the transform of g is $\exp(i2\pi cu) \times F(u)$, and in Eq. (5) [or (6)] the exponential factors in the expression for $A_{k,g}$ ($\ddot{A}_{k,g}$) cancel one another, leaving $A_{k,g} = A_{k,f}$ ($\ddot{A}_{k,g} = \ddot{A}_{k,f}$). (iii) Suppose that $a_{k,g} = a_{k,f}$. Then, from Eq. (5),

$$G\left(-\sum_{j=1}^{k} u_j\right)\prod_{j=1}^{k} G(u_j) = F\left(-\sum_{j=1}^{k} u_j\right)\prod_{j=1}^{k} F(u_j). \tag{7}$$

If h is any third function, the transforms of the convolutions $h * g$ and $h * f$ are $H(u)G(u)$ and $H(u)F(u)$, respectively, and if Eq. (7) holds for G and F, it also holds for HG and HF. The same argument holds for periodic images through Eq. (6). Property (iv) follows from the same reasoning applied to FF and GG. Property (v) follows from the fact that, if Eq. (7) holds for the transforms of $f(x)$ and $g(x)$, it also holds for the transforms of $f(ax)$ and $g(ax)$, i.e., for $(1/|a|)F(u/a)$ and $(1/|a|)G(u/a)$.

The following symmetry properties of the bispectrum are also immediate consequences of Eqs. (5) and (6):

(vi)
$$A_{2,f}(u,v) = A_{2,f}(v,u),$$
$$\ddot{A}_{2,f}(u,v) = \ddot{A}_{2,f}(v,u);$$

(vii)
$$A_{2,f}(u,v) = A_{2,f}(u,-u-v),$$
$$\ddot{A}_{2,f}(u,v) = \ddot{A}_{2,f}(u,-u-v);$$

(viii)
$$A_{2,f}(u,v) = A_{2,f}^{*}(-u,-v),$$
$$\ddot{A}_{2,f}(u,v) = \ddot{A}_{2,f}^{*}(-u,-v).$$

[The asterisks in relations (viii) denote complex conjugation.] Relations (vi)–(viii) imply that the bispectrum $A_{2,f}(u,v)$ [$\ddot{A}_{2,f}(u,v)$] is determined by its values in any one octant, e.g., the octant $0 \le u$, $0 \le v \le u$.

D. Uniqueness of Higher-Order ACF's of Integrable Images

1. *Uniqueness for Images of Finite Size*

We begin by giving a new proof of the triple-correlation-uniqueness theorem of Bartelt *et al.*[11] for images with bounded support:

Theorem 1. Suppose that f is an integrable image function and that, for some $b \ge 0$, $f(x) \equiv 0$ for $|x| \ge b$. Then $a_{2,g} = a_{2,f}$ for another image g if and only if $g(x) = f(x + c)$ for some constant c.

Proof: "If" follows from relation (ii) in Subsection 2.C. To prove "only if," we start with the fact that if $a_{2,g} = a_{2,f}$ then $A_{2,g} = A_{2,f}$, so, from Eq. (5),

$$G(u)G(v)G(-u - v) = F(u)F(v)F(-u - v) \tag{8}$$

for all u and v. We set $u = v = 0$ in Eq. (8); then $G(0)^3 = F(0)^3$, so $F(0) = G(0)$. If $F(0) = 0$, then $f = 0$, since $F(0) = \int_{-\infty}^{\infty} f(x)\mathrm{d}x$, and in that case $g = 0$ as well. So we assume $F(0)$ and $G(0)$ are not zero, and without loss of generality we can assume that their common value is 1. [If it is not, we can divide $a_{2,f}$ by $F(0)^3$ and $a_{2,g}$ by $G(0)^3$ and show that $g(x)/G(0) = f(x + c)/F(0)$, i.e., $g(x) = f(x + c)$, since $G(0) = F(0)$.] Then f and g are probability-density functions, and Eq. (8) is a relationship between the characteristic functions of these densities {i.e., the expectations $E[\exp(-i2\pi u\mathbf{x}_f)]$ and $E[\exp(-i2\pi u\mathbf{x}_g)]$, where the random variables $\mathbf{x}_f$ and $\mathbf{x}_g$ have densities f and g}. We use two well-known properties of characteristic functions (e.g., see Feller,[21] Chap. XV): (i) Every characteristic function is continuous, and (ii) every probability-density function with bounded support is uniquely determined by the values of its characteristic function in a neighborhood of the origin. Since $F(0) = G(0) = 1$, property (i) implies that there is an interval around the origin, say $(-\beta,\beta)$, in which both F and G are nonvanishing. Now write F and G in exponential form: $F(u) = |F(u)|\exp[iPhaF(u)]$ and $G(u) = |G(u)|\exp[iPhaG(u)]$. Since $F(u)$ and $G(u)$ are nonvanishing and continuous for $u \in (-\beta,\beta)$, $PhaG(u)$ and $PhaF(u)$ are defined and continuous in that interval, and both equal 0 at $u = 0$. Setting $v = -u$ in Eq. (8) shows that $|G(u)| = |F(u)|$, so if we can prove that in some neighborhood of the origin $PhaG(u) = PhaF(u) + 2\pi cu$ for some constant c, then $G(u) = \exp(i2\pi cu)F(u)$ in that neighborhood. From Eq. (8) we have

$$PhaG(u) + PhaG(v) - PhaG(u + v)$$
$$= PhaF(u) + PhaF(v) - PhaF(u + v) + 2\pi N(u,v) \tag{9}$$

for all values of u,v and $u + v$ for which F and G are not zero, with $N(u,v)$ an integer. For every fixed $v' \in (-\beta/2,\beta/2)$, $N(u,v')$ is a linear combination of constants and continuous functions of u and hence is itself continuous, and $N(0,v') = 0$. Since $N(u,v')$ is an integer, it must be zero for all $u \in (-\beta/2,\beta/2)$, so $N(u,v) \equiv 0$ for all $u,v \in (-\beta/2,\beta/2)$. Let $D(u) = PhaG(u) - PhaF(u)$; $D(u)$ is continuous for $u \in (-\beta/2,\beta/2)$. Rearranging Eq. (9) with $N(u,v) = 0$, we have for all u,v in the interval $(-\beta/2,\beta/2)$

$$D(u + v) = D(u) + D(v). \tag{10}$$

Equation (10) is the classic Cauchy functional equation. Aczel[22] shows that, if D is continuous and Eq. (10) holds for all u,v in any interval containing the origin, then over that interval $D(u) = bu$ for some constant b. Setting $b = 2\pi c$, we have $PhaG(u) = PhaF(u) + 2\pi cu$, so $G(u) = \exp(i2\pi cu)F(u)$ in a neighborhood of the origin. Consequently, in that neighborhood the characteristic function of the probability-density function $g(x)$ agrees with that of some density $f(x + c)$. Since $f(x)$ has bounded support, $f(x + c)$ does also, and thus its characteristic function is completely determined by its values in a neighborhood of the origin. Thus $G(u) = \exp(i2\pi cu)F(u)$ for all u, so $g(x) = f(x + c)$, and Theorem 1 is proved.

2. *Reconstruction Algorithms for Finite Images*

Theorem 1 guarantees that every finite-sized 1-D image f is uniquely determined up to the form $f(x + c)$ by its triple correlation but does not show explicitly how the family $\{f(x + c)\}$ can be recovered from $a_{2,f}$. We are aware of two basic approaches to this problem. One involves a recursive reconstruction of the Fourier transform F from the bispectrum $A_{2,f}$. The other uses the derivatives of the bispectrum (or, equivalently, certain moments of the triple correlation) to recover the moments of f, which determine the power-series expression for its transform. We discuss first the recursive approach and then the one based on moments.

Equation (5) shows that the amplitude spectrum $|F(u)|$ can be obtained immediately from $A_{2,f}(u,v)$ through the relationship

$$|F(u)| = \{A_{2,f}(u,0)/[A_{2,f}(0,0)]^{1/3}\}^{1/2}. \quad (11)$$

However, there is no analogous direct expression relating the phase spectrum $PhaF(u)$ to $A_{2,f}(u,v)$. To recover $PhaF(u)$, several authors[11,13–15] have proposed recursive algorithms based on the fact that Eq. (5) implies that the phase of F is related to the phase of $A_{2,f}$ by

$$PhaF(u+v) = PhaF(u) + PhaF(v) - PhaA_{2,f}(u,v) \quad (12)$$

for all u,v, $u+v$ at which F does not vanish. Since $A_{2,f}$ determines F only up to the form $\exp(i2\pi cu)F(u)$, we can assign one frequency an arbitrary value, say, $PhaF(u_1) = \Theta$ ($\Theta = 0$ being the natural choice). Then to recover $PhaF$ at frequencies that are multiples of u_1 one could try to use the simple recursion

$$PhaF(nu_1) = PhaF[(n-1)u_1] + \Theta - PhaA_{2,f}[(n-1)u_1, u_1]. \quad (13)$$

For example, if f is assumed to vanish outside a finite interval $(-w/2, w/2)$, one might take u_1 to be $1/w$, since in that case the sampling theorem implies that F is determined by its values on the set $\{n/w: n = 0, \pm 1, \pm 2, \ldots\}$. If F is nonvanishing on all multiples of u_1, recursion (13) will automatically deliver all the phases $PhaF(nu_1)$. However, if F vanishes on the sequence $u_1, 2u_1, \ldots$, recursion (13) cannot be used beyond the first n for which $F(nu_1) = 0$: at that point $PhaF(nu_1)$ is undetermined, and $PhaF[(n+1)u_1]$ cannot be calculated from it. In this case $Pha[(n+1)u_1]$ may still be recoverable by means of Eq. (12) by using some combination $Pha(ku_1)$ and $Pha[(n+1-k)u_1]$ with $k < n$: the bispectrum must be examined to determine whether this is possible. Failing that, a finer sampling lattice can be tried. In general one cannot specify in advance a frequency u_1 for which recursion (13) is guaranteed to succeed: one needs first to identify the zeros of F and then to find a u_1 for which $F(nu_1)$ is never zero and $1/u_1$ is an adequate sampling rate for F. Bartelt *et al.*[11] note that, in principle, a value of u_1 that will work for all frequencies nu_1 up to any desired limit can always be found, because the Fourier transform of a function with bounded support can have only a finite number of zeros in any finite interval. However, in practice, phase recovery for finite portions of images whose infinite versions are not uniquely determined by their triple correlations (e.g., those in Fig. 1) will become increasingly difficult as the size of the observation window grows and the bispectrum of the windowed portion converges to that of the infinite version. (This point is discussed below in Subsection 2.F.)

An alternative approach to image reconstruction can be based on the fact that the derivatives of the bispectrum $A_{2,f}(u,v)$ along the line $v = u$ determine the moments of f [more precisely, the moments of a certain member of the family $\{f(x+c)\}$], which in turn can be used to reconstruct the transform $F(u)$ [up to a factor $\exp(i2\pi cu)$] and thus f itself. This approach may be too fragile for practical use, but the argument has mathematical interest because it provides a constructive proof of Theorem 1. To simplify matters, assume that $\int_{-\infty}^{\infty} f(x) = 1$ [i.e., $F(0) = A_{2,f}(0,0)^{1/3} = 1$; if not, we can begin by dividing $A_{2,f}$ by $A_{2,f}(0,0)$ and recover $f(x+c)/F(0)$]. Let $f_0(x)$ be the unique member of the family $\{f(x+c)\}$ for which $\int_{-\infty}^{\infty} xf(x+c)\mathrm{d}x = 0$. Since f_0 has finite support, its transform F_0 can be expressed as a Taylor series about the origin, which is valid for all u:

$$F_0(u) = \int_{-\infty}^{\infty} \exp(-i2ux)f_0(x)\mathrm{d}x = \sum_{n=0}^{\infty} (-i2\pi u)^n \mu_n/n!, \quad (14)$$

where μ_n is the nth moment of f_0: $\mu_n = \int_{-\infty}^{\infty} x^n f_0(x)\mathrm{d}x$. If the moments μ_n can be determined, then F_0 can be constructed from Eq. (14) and f_0 can be recovered by Fourier inversion. We know that $\mu_0 = 1$, and $\mu_1 = 0$ by construction. To obtain the other moments, let $Q(u) = A_{2,f}(u,u)$. From Eq. (5), $Q(u) = F(u)^2 F(-2u)$, so Q is the characteristic function of a random variable $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 - 2\mathbf{x}_0$, where $\mathbf{x}_j$, $j = 0,1,2$, are independent random variables with common density f_0; i.e., $Q(u) = E\{\exp[-i2\pi u(\mathbf{x}_1 + \mathbf{x}_2 - 2\mathbf{x}_0)]\}$. Q is nonvanishing in a neighborhood of the origin, so $\log Q$ is defined in that neighborhood. Moreover, Q is the characteristic function of a random variable that has finite absolute moments of all orders (since the density of $\mathbf{y}$ has bounded support), so Q is infinitely differentiable at the origin, and thus $\log Q$ is also. Now let $L_Q(u) = \log Q(u/-2\pi)$ and $L_F(u) = \log F_0(-2\pi u)$. Then

$$L_Q(u) = 2L_F(u) + L_F(-2u),$$

and differentiating this equation n times yields

$$L_Q{}^{(n)}(u) = 2L_F{}^{(n)}(u) + (-2)^n L_F{}^{(n)}(-2u).$$

Thus at $u = 0$

$$i^{-n}L_F{}^{(n)}(0) = i^{-n}L_Q{}^{(n)}(0)[2 + (-2)^n]^{-1} \quad (15)$$

for $n = 2,3,\ldots$. The quantities $\kappa_n = i^{-n}L_F{}^{(n)}(0)$ on the left-hand side of Eq. (15) are the cumulants of the unknown density function f_0 whose moments we seek, and Eq. (15) shows that, for $n \geq 2$, κ_n can be obtained from the known derivative $L_Q{}^{(n)}(0)$. It is a fact of probability theory that the moments μ_n of a density function can be calculated directly from its cumulants: $\mu_1 = \kappa_1, \mu_2 = \kappa_2, \mu_3 = \kappa_3 + 3\kappa_1\kappa_2 + (\kappa_3)^3, \ldots$. (Lukacs[23] gives a general formula for calculating the moments of a density from its cumulants.) Here μ_1 is 0 by construction, and Eq. (15) shows that all the remaining moments of f_0 can be obtained from the successive derivatives of $\log A_{2,f}(-u/2\pi, -u/2\pi)$. Consequently, F_0 and thus f_0 can be reconstructed from $A_{2,f}$, as is claimed. The argument provides a direct proof of Theorem 1 and shows in addition that, when f has finite support, all the information in its triple-correlation function is carried by a countable set of values: the derivatives of $A_{2,f}(u,u)$ at $u = 0$ or, equivalently, the triple-correlation moments that correspond to those derivatives:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (-i2\pi)^n (s_1 + s_2)^n a_{2,f}(s_1,s_2)\mathrm{d}s_1\mathrm{d}s_2 = A_{2,f}{}^{(n)}(0,0).$$

3. Higher-Order Autocorrelations of Infinite Images: Examples of Nonuniqueness

Theorem 1 shows that, for integrable images of finite size, the information available from the entire set of higher-order ACF's of an image is already fully contained in its triple correlation: $a_{2,f}$ determines f up to the form $f(x + c)$, and, in view of relation (ii) in Subsection 2.C, the unidentified constant c cannot be obtained from any higher-order ACF. For images with infinite support the uniqueness situation is not so straightforward. While Theorem 1 can be extended to certain useful classes of infinite images, as is shown below by Theorems 3–5, in general it is not the case that all integrable images with infinite support are uniquely determined up to translation by their triple-correlation functions. In fact, no ACF of any fixed order uniquely determines all the integrable images up to translation: for every k there are images f and g for which $a_{k,g} = a_{k,f}$ but $g(x) \neq f(x + c)$ for any c.

This point is demonstrated by the following example: For any integer $k \geq 2$ the integrable images

$$f(x) = \mathrm{sinc}^2(x)[1 + \cos 2\pi(k + 1)x], \tag{16}$$

$$g(x) = \mathrm{sinc}^2(x)[1 + \sin 2\pi(k + 1)x] \tag{17}$$

have the same kth-order ACF and clearly are not translations of each other: f is symmetric; g is not. (Figure 1 illustrates f and g for $k = 2$.) The fact that $a_{k,f} = a_{k,g}$ for Eqs. (16) and (17) follows from the fact that the transforms of f and g satisfy the relationship

$$F\left(-\sum_{j=1}^{k} u_j\right)\prod_{j=1}^{k} F(u_j) = G\left(-\sum_{j=1}^{k} u_j\right)\prod_{j=1}^{k} G(u_j) \tag{18}$$

for all arguments $(u_1, \ldots, u_k)$, and thus, from Eq. (5), $A_{k,f} = A_{k,g}$.

The transforms of Eqs. (16) and (17) are

$$F(u) = \Lambda(u) + (1/2)\Lambda[u - (k + 1)] + (1/2)\Lambda[u + (k + 1)], \tag{19}$$

$$G(u) = \Lambda(u) - (i/2)\Lambda[u - (k + 1)] + (i/2)\Lambda[u + (k + 1)], \tag{20}$$

where Λ is the triangle function: $\Lambda(u) = 1 - |u|$ for $|u| \leq 1$ and 0 elsewhere. Thus f and g are band-limited functions, and since $F(0) = G(0) = 1$, both are integrable. Rockwell and Yellott[7] show that transforms (19) and (20) satisfy Eq. (18) for any $k \geq 2$, and thus $A_{k,f} = A_{k,g}$, so Eqs. (16) and (17) have the same kth-order ACF. [For any fixed k, Eq. (18) fails for $k + 1$, so Eqs. (16) and (17) do not have identical $(k + 1)$st-order ACF's.] The proof of this for an arbitrary k is straightforward but too long to reproduce here in full. To illustrate the argument we prove the special case $k = 2$, showing that the images

$$f(x) = \mathrm{sinc}^2(x)(1 + \cos 2\pi 3x), \tag{21}$$

$$g(x) = \mathrm{sinc}^2(x)(1 + \sin 2\pi 3x) \tag{22}$$

have the same triple correlation. In this case Eq. (18) holds if

$$F(u)F(v)F(-u - v) = G(u)G(v)G(-u - v) \tag{23}$$

for all u, v, with

$$F(u) = \Lambda(u) + (1/2)\Lambda(u - 3) + (1/2)\Lambda(u + 3),$$

$$G(u) = \Lambda(u) - (i/2)\Lambda(u - 3) + (i/2)\Lambda(u + 3).$$

Let I_0 denote the interval $(-1, 1)$, I_3 the interval $(2, 4)$, I_{-3} the interval $(-4, -2)$, and C the complement of the union $I_0 \cup I_3 \cup I_{-3}$. Then F and G vanish on C; $G(u) = F(u)$ for $u \in I_0$, $G(u) = -iF(u)$ for $u \in I_3$, and $G(u) = iF(u)$ for $u \in I_{-3}$. If either $u \in C$ or $v \in C$, both sides of Eq. (23) vanish, so the identity holds. If both $u \in I_0$ *and* $v \in I_0$, then $-u - v$ is either in C and both sides of Eq. (23) vanish or $-u - v \in I_0$ and both sides are equal factor by factor. If $u \in I_0$ and $v \in I_3$, then $G(u) = F(u)$, $G(v) = -iF(v)$, and $-5 < -u - v < -1$, so either $-u - v \in C$ and both sides of Eq. (23) vanish or $-u - v \in I_{-3}$ and $G(-u - v) = +iF(-u - v)$, in which case the $+i$ and $-i$ factors on the G side of Eq. (23) cancel each other and the equality holds. Similarly, if $u \in I_0$ and $v \in I_{-3}$, then $-u - v$ is either in C and both sides of Eq. (23) vanish or in I_3 and the $\pm i$ factors cancel. If u and v are both in I_3 or are both in I_{-3}, then $-u - v$ is in C. Finally, if $u \in I_3$ and $v \in I_{-3}$, $G(u)G(v) = F(u)F(v)$, and $-u - v$ is either in I_0 and $G(-u - v) = F(-u - v)$ or in C, so Eq. (23) holds in either case.

[The same argument can be used to show that Eqs. (21) and (22) have the same triple correlation when $\sin 2\pi 3x$ in Eq. (22) is replaced by $-\sin 2\pi 3x$ or by $-\cos 2\pi 3x$. A geometrical understanding of the lack of uniqueness for these images can be gained from Fig. 5 in Subsection 2.F, which illustrates the common support of their bispectra.]

The proof that Eqs. (16) and (17) satisfy Eq. (18) for an arbitrary $k \geq 2$ involves the same sort of argument checking on a more elaborate scale.[7]

When f and g have the same kth-order ACF, the same is also true of $h * f$ and $h * g$ for any third function h [property (iii) in Subsection 2.C] and of $f * f$ and $g * g$ [property (iv)], so functions (16) and (17) can be used to construct an infinite variety of pairs of distinct band-limited images f and g with identical kth-order ACF's. And since rescaling such images will not alter the identity between their kth-order ACF's [property (v)], we can construct examples of distinct band-limited functions with an arbitrarily small bandwidth that have identical kth-order ACF's. In view of such examples, the only completely general uniqueness theorem that one can prove for integrable images is the following:

Theorem 2. If f and g are integrable images and $a_{k,g} = a_{k,f}$ for all k, then $g(x) = f(x + c)$.

Proof: The proof of Theorem 1 shows that if $a_{2,g} = a_{2,f}$ there is a neighborhood Δ of the origin in which the transforms G and F are nonvanishing and $G(u) = \exp(i2\pi cu) \times F(u)$. Since $a_{k,g} = a_{k,f}$ for all k, Eq. (5) implies that $F(-ku)F(u)^k = G(-ku)G(u)^k$ for all u and k. For any $v \in \mathbf{R} - \Delta$ there is a $u \in \Delta$ such that $v = -ku$ for some k, and since $G(u) = \exp(i2\pi cu)F(u) \neq 0$ for all $u \in \Delta$, we have

$$G(v)[G(u)]^k = G(v)[\exp(i2\pi cu)F(u)]^k = F(v)[F(u)]^k.$$

Dividing both sides of the second equality by $[\exp(i2\pi cu) \times F(u)]^k$ (with $ku = -v$) yields $G(v) = \exp(i2\pi cv)F(v)$. Since this holds for all v, $g(x) = f(x + c)$.

4. *Uniqueness Theorems for Special Classes of Infinite Images*

The examples given in Subsection 2.D.3 show that, in general, the triple-correlation function of an integrable image with infinite support does not uniquely determine that image up to translation. However, it is possible to extend Theorem 1 to certain natural classes of infinite images. One approach is to show that the relationship $PhaG(u) = PhaF(u) + 2\pi cu$, which holds in some neighborhood of the origin for any images f and g that satisfy $a_{2,g} = a_{2,f}$, can be extended by using Eq. (9) to any point u at which F and G are nonvanishing, provided that F has at most a finite number of zeros below u. The following theorem summarizes the results of this approach.

Theorem 3. If f is an integrable image and $a_{2,g} = a_{2,f}$ for another image g, then $g(x) = f(x + c)$ for some constant c if the Fourier transform of f satisfies any one of the following conditions:

(a) $F(u)$ is nonvanishing for all u.

(b) $F(u) \equiv 0$ for $|u| \geq$ some $b > 0$, and $F(u)$ is nonvanishing for $|u| < b$.

(c) $F(u) \equiv 0$ for $|u| \geq$ some $b > 0$, and $F(u) = 0$ for at most a finite number of values in the interval $(-b, b)$.

(d) $F(u) = 0$ for at most a finite number of values in every finite subinterval of the real line.

[Obviously, condition (c) includes condition (b) as a special case, and likewise (d) includes (a), but we state them separately for clarity.]

Proof: See Appendix A.

Examples of the various cases were mentioned in Section 1. We note that, if a nonzero image has bounded support, its transform can have only a finite number of zeros in any finite interval, since in this case its complex transform is an entire analytic function (e.g., see Lukacs,[23] Theorem 7.2.3) and cannot vanish infinitely often on any finite interval of the real axis without vanishing everywhere. Thus condition (c) implies that when, e.g., $\mathrm{sinc}^2(x)$ is convolved with any finite image, the resulting image is determined up to translation by its triple correlation. Part (d) shows that the same is true of finite images convolved with impulse responses whose transforms are nonvanishing, e.g., Gaussians.

Theorem 3 excludes images whose spectra contain nonzero regions separated by an interval of zeros. Images (21) and (22), which are not determined by their triple correlations, have transforms of that sort, but the fact that the spectrum of an image contains intervals of zeros does not necessarily mean that the image is undetermined by its triple correlation. The critical factor is the size of those intervals—in particular, their size relative to the width of the neighborhood of the origin in which the transform is nonvanishing. Let Δ_f be that neighborhood for an image f. We say that a point u, at which the transform $F(u) \neq 0$, can be finitely linked to Δ_f if there is a point p_0 in Δ_f and a sequence of numbers $\alpha_1, \ldots, \alpha_n$, $0 < \alpha_i \leq 1$, such that

$$u = \left(1 + \sum_{i=1}^{n} \alpha_i\right) p_0 \tag{24}$$

and at each point $p_i = (1 + \alpha_1 + \cdots + \alpha_i)p_0$, $F(p_i) \neq 0$. (The points p_i, $0 < i < n$ serve as stepping stones linking u to p_0.)

Theorem 4. If f is an image for which every point u at which the transform $F(u) \neq 0$ can be finitely linked to Δ_f and if g is another image with $a_{2,g} = a_{2,f}$, then $g(x) = f(x + c)$ for some constant c.

As an example, $f(x) = \mathrm{sinc}^2(x)(1 + \cos 5\pi x)$ has the transform $F(u) = \Lambda(u) + (1/2)\Lambda(u - 2.5) + (1/2)\Lambda(u + 2.5)$. Points in the interval $(1.5, 3.5)$ can obviously be linked to the central interval $(-1, 1)$, so f is determined by its triple correlation.

Proof: Suppose that $F(u) \neq 0$, and let p_i, $i = 1, \ldots, n$ be the sequence linking u to a point p_0 in the neighborhood Δ_f by means of Eq. (24). The proof of condition (b) of Theorem 3 shows that, for some constant c (independent of u), $G(v) = \exp(i2cv)F(v)$ for all v in Δ_f, in particular at $v = -p_0$ and $-\alpha_1 p_0$. Applying Eq. (8), we have $G(p_1) = \exp(i2\pi cp_1)F(p_1)$, which (from $|G| = |F|$ and the oddness of the phase) implies also that $G(-p_1) = \exp(-i2\pi cp_1) \times F(-p_1)$. Then, if $G(p_j) = \exp(i2\pi cp_j)F(p_j)$, we have $G(-\alpha_{j+1}p_0) = \exp(-i2\pi c\alpha_{j+1}p_0)F(-\alpha_{j+1}p_0)$ and $G(-p_j) = \exp(-i2\pi cp_j)F(p_j)$, and applying Eq. (8) gives $G(p_{j+1}) = \exp(i2\pi cp_{j+1})F(p_{j+1})$. Proceeding step by step from p_1, we eventually obtain $G(u) = \exp(i2\pi cu)F(u)$, so $g(x) = f(x + c)$.

The proof of Theorem 1 shows that whenever f and g have the same triple correlation there must be a neighborhood of the origin in which $G(u) = \exp(i2\pi cu)F(u)$, and Theorems 3 and 4 broaden the scope of triple-correlation uniqueness by showing that this equality can be extended to all u if the zeros of F satisfy certain conditions. Another way to generalize Theorem 1 is to note that its proof used the assumption that f has bounded support only to guarantee that, for every c, $f(x + c)$ is uniquely determined by the derivatives of its transform $\exp(i2\pi cu)F(u)$ at $u = 0$, which agree with those of $G(u)$, implying $g(x) = f(x + c)$. Bounded support is a sufficient condition for this but not a necessary one. If they exist, the derivatives of $\exp(i2\pi cu)F(u)$ are the sequence $\{(-i2\pi)^n \mu_n\}$, where μ_n is the nth moment of $f(x + c)$, i.e., $\mu_n = \int_{-\infty}^{\infty} x^n f(x + c)\mathrm{d}x$. When $f(x)$ and thus $f(x + c)$ have bounded support, the moment sequence $\{\mu_n\}$ uniquely determines a power-series expression for the transform $\exp(i2\pi cu)F(u)$ and thus identifies $f(x + c)$ itself. However, a function may be uniquely determined by its moments even though its transform does not have a power-series representation that is valid for all u. [This is true, for example, of the one-sided exponential density $f(x) = \exp(-x), x \geq 0$, as is discussed below.] Clearly if, for every c, $f(x + c)$ belongs to a class of functions that are uniquely determined by their moments, then f is determined up to translation by its triple correlation. And if $f(x)$ itself is determined by its moments, so too is $f(x + c)$ for every c. [This follows from the fact that there is a one-to-one correspondence between the moments $\{\mu_n\}$ of f and its cumulants $\{\kappa_n\}$. If f is uniquely determined by its moments, the same is true of its cumulants. The cumulants of $f(x + c)$ are $\{\kappa_1 + c, \kappa_2, \kappa_2, \ldots\}$, so when f is determined by its cumulants, $f(x + c)$ is also.] Thus whenever f is uniquely determined by its moments, it is also uniquely determined up to translation by its triple correlation.

The problem of characterizing classes of functions that are uniquely determined by their moments has a large literature that we will not attempt to summarize; Shohat and Tamarkin[17] and Akhiezer[18] provide reviews. Instead we describe a single condition involving moments that is sufficient to guarantee that an image is determined by its triple correlation. Characteristic-function theory[23] shows that, if all the moments of an image f are finite and satisfy the condition

$$\lim_{n\to\infty} \sup(|\mu_n|/n!)^{1/n} = \lambda < \infty, \tag{25}$$

then $F(u)$ has a unique extension to a function $F(z)$, with z complex, which is analytic in an open disk $|z| < 1/2\pi\lambda$. In this case $F(z)$ is analytic for all z in an open horizontal strip containing the real axis and is determined by analytic continuation for all real z by the derivatives $F^{(n)}(0)$. Since $\exp(i2\pi cz)$ is an entire function for any constant c, $\exp(i2\pi cz)F(z)$ will be analytic in the same strip and thus be determined for all real z by its derivatives at zero. So we have the following:

Theorem 5. If f is an integrable image that is uniquely determined by its moments [for example, if its moments satisfy Eq. (25)] and g is another image for which $a_{2,g} = a_{2,f}$, then $g(x) = f(x + c)$.

[We note that an image may be uniquely determined by its moments without satisfying Eq. (25): Shohat and Tamarkin[17] give the example $f(x) = \exp(-x^{1/2})$, $x \geq 0$. Uniqueness in that case is established by Carleman's theorem: An image is determined by its moments if $\sum_{n=1}^{\infty} \mu_{2n}{}^{-1/2n} = \infty$. Shohat and Tamarkin also show that, if the exponent 1/2 in this example is replaced with any positive value less than 1/2, f is no longer determined by its moments, although all the moments are finite.]

If f and g are two images whose individual transforms F and G satisfy Eq. (25) and thus are both analytic in some neighborhood of the real axis, the product FG is also analytic in such a neighborhood, hence determined by its derivatives at 0, so $f*g$ is determined up to translation by its triple correlation. The Gaussian $f(x) = \exp(-ax^2)$ satisfies condition (25), so we have another proof that any finite image convolved with a Gaussian impulse response is determined by its triple correlation.

It was shown in Subsection 2.D.2 that all the moments of a finite image f can be recovered from the derivatives of $A_{2,f}(u,u)$ at $u = 0$. The same procedure will work for any image whose moments satisfy Eq. (25), so for all such images it is the case that all the information in the bispectrum $A_{2,f}(u,v)$ is carried by its values along the line $v = u$ in a neighborhood of $u = 0$. However, in general this does not mean that one can directly reconstruct f by inverting a power-series expression for its Fourier transform of the form of Eq. (14), as is always possible for finite images. The transforms of finite images are always entire functions whose power series converge for all arguments [as is shown by the fact that λ in Eq. (25) is zero in this case]. But an image can satisfy Eq. (25) and thus be uniquely determined by its moments without its transform's being entire. The one-sided exponential function $f(x) = \exp(-x)$ for $x \geq 0$ is an example: here $\mu_n = n!$, so $\lambda = 1$, and the power series of its transform, $F(u) = (1 - i2\pi u)/(1 + 4\pi^2 u^2)$, converges only for $|u| < 1/2\pi$. In this case the values of $F(u)$ for $|u| \geq 1/2\pi$ are still determined in principle by its derivatives at $u = 0$ (i.e., by the moments μ_n), but their actual calculation by analytic continuation would not be straightforward. Thus Theorem 5 guarantees that images whose moments satisfy Eq. (25) are uniquely determined up to translation by their triple correlations but does not provide a practical way of reconstructing all such images from those moments.

E. Uniqueness of Higher-Order Autocorrelations of Periodic 1-D Images

By comparison with the rich subset of integrable images that are determined up to translation by their triple correlations, the uniqueness properties of the higher-order ACF's of infinitely extended periodic images seem rather bleak. In general, for every integer n, there are pairs of band-limited periodic images that are not translations of each other but whose ACF's agree for all orders up through n, and there do not seem to be many interesting subclasses that escape this ambiguity. Klein and Tyler[4] give the following example demonstrating the limited possibilities here: For every integer $k \geq 2$, the functions

$$f(x) = 2 + \cos 2\pi x + \cos 2\pi kx, \tag{26}$$

$$g(x) = 2 + \cos 2\pi x - \cos 2\pi kx \tag{27}$$

have the same $(k - 1)$st-order ACF. (Figure 2 illustrates the case $k = 3$, where f and g have the same triple correlation.) To prove this, note that the Fourier transforms of the single-period ($p = 1$) segments of Eqs. (26) and (27) are, respectively,

$$F_1(u) = 2\,\mathrm{sinc}(u) + (1/2)[\mathrm{sinc}(u-1) + \mathrm{sinc}(u+1)] + (1/2)[\mathrm{sinc}(u-k) + \mathrm{sinc}(u+k)], \tag{28}$$

$$G_1(u) = 2\,\mathrm{sinc}(u) + (1/2)[\mathrm{sinc}(u-1) + \mathrm{sinc}(u+1)] - (1/2)[\mathrm{sinc}(u-k) + \mathrm{sinc}(u+k)]. \tag{29}$$

It follows from Eq. (4) that $\ddot{A}_{k-1,f} = \ddot{A}_{k-1,g}$, and thus $\ddot{a}_{k-1,f} = \ddot{a}_{k-1,g}$, if

$$F_1\left(-\sum_{j=1}^{k-1} m_j\right)\prod_{j=1}^{k-1} F_1(m_j) = G_1\left(-\sum_{j=1}^{k-1} m_j\right)\prod_{j=1}^{k-1} G_1(m_j) \tag{30}$$

for all sets of integers $m_1, \ldots, m_{k-1}$. For integer values of u, F_1 and G_1 vanish except at $u = 0, \pm 1$, and $\pm k$, and for those arguments F_1 and G_1 differ only at $\pm k$: $G_1(\pm k) = -F_1(\pm k)$. Thus both sides of Eq. (30) vanish and the identity is always true if $m_1, \ldots, m_{k-1}$ do not all come from the set $0, \pm 1, \pm k$. For Eq. (30) to fail, then, the set of k arguments $m_1, \ldots, m_{k-1}, -\sum_{j=1}^{k-1} m_j$ must contain an odd number of occurrences of $\pm k$, and $-\sum_{j=1}^{k-1} m_j$ must be 0, ± 1, or $\pm k$. When that sum is 0, $+k$ and $-k$ must occur equally often in the set $m_1, \ldots, m_{k-1}$, since an extra k cannot be canceled by the sum of at most $k - 2$ occurrences of -1 or $+1$. Consequently, in this case the total number of occurrences of $+k$ and $-k$ in the set $m_1, \ldots, m_{k-1}$, $-\sum_{j=1}^{k-1} m_j$ must be even, and Eq. (30) is true. The same sort of argument can be made for the cases $\sum_{j=1}^{k-1} m_j = \pm 1$ and $\pm k$, showing that there can never be an odd number of occurrences of $\pm k$ in the set $m_1, \ldots, m_{1-1}, -\sum_{j=1}^{k-1} m_j$, so Eq. (30) is always true, and $\ddot{A}_{k-1,f} = \ddot{A}_{k-1,g}$. [However, $\ddot{A}_{k,f} \neq \ddot{A}_{k,g}$, so Eqs. (26) and (27) have different kth-order ACF's. To show this, substitute k for $k - 1$ in Eq. (30) and set $m_1 = m_2 = \ldots = m_k = 1$. Then $-\sum_{j=1}^{k} m_j = k$, making the left-hand side of Eq. (30) positive while the right-hand side is negative.]

Since an equality between the kth-order ACF's of f and g cannot be undone by filtering [property (iii) in

Subsection 2.C] or rescaling [property (v)], the fact that $\ddot{a}_{k-1,f} = \ddot{a}_{k-1,g}$ for Eqs. (26) and (27) shows that the same is true for any images of the forms

$$f(x) = L(2 + C_1 \cos 2\pi\phi x + C_2 \cos 2\pi k\phi x),$$

$$g(x) = L(2 + C_1 \cos 2\pi\phi x - C_2 \cos 2\pi k\phi x),$$

with $k \geq 2$, $L > 0$ and $\phi > 0$, and $0 < C_1, C_2 \leq 1$. Property (iv) in Subsection 2.C shows that the pair f^*f and g^*g constructed from any of these functions will also have identical ACF's of order $k - 1$.

In view of these examples, the only completely general uniqueness theorem that one can hope to prove here is that every periodic image is uniquely determined up to translation if all its ACF's are known. We showed earlier (Theorem 2) by a simple argument that this is true of all integrable images, but the periodic case is more difficult to decide one way or the other. We leave it as an open problem[19] and prove instead an easy weaker result: Every periodic image with period p is determined up to translation by its ACF's of all orders if its spectrum is nonvanishing at the fundamental frequency $1/p$:

Theorem 6. If f is a periodic image with period p and $F(1/p) \neq 0$, then $\ddot{a}_{k,g} = \ddot{a}_{k,f}$ for all k for another image g if and only if $g(x) = f(x + c)$.

Proof: "If" is property (ii) in Subsection 2.C. To show "only if," we start with the fact that $\ddot{a}_{2,g} = a_{2,f}$ implies, from Eq. (6), that $|G(n/p)| = |F(n/p)|$ for all n. Since $F(1/p) \neq 0$, $F(-1/p) \neq 0$ and $G(-1/p) \neq 0$, and we can write

$$\begin{aligned} G(-1/p) &= |F(-1/p)| \exp\{i[PhaG(-1/p) - PhaF(-1/p)]\} \\ &\quad \times \exp[iPhaF(-1/p)] \\ &= F(-1/p)\exp(-i2\pi c/p), \end{aligned}$$

where $-2\pi c/p = PhaG(-1/p) - PhaF(-1/p)$. So

$$F(-1/p)/G(-1/p) = \exp(i2\pi c/p).$$

Then $\ddot{a}_{k,g} = \ddot{a}_{k,f}$ for any k implies, from Eq. (6), that

$$[G(-1/p)]^k G(k/p) = [F(-1/p)]^k F(k/p),$$

and division yields $G(k/p) = \exp(i2\pi k/p)F(k/p)$. So $G(u) = \exp(i2\pi cu)F(u)$ for all u, and $g(x) = f(x + c)$.

The next result is a periodic analog to conditions (a) and (b) of Theorem 3:

Theorem 7. If f is a periodic image with period p and $\ddot{a}_{2,g} = \ddot{a}_{2,f}$ for another image g, then $g(x) = f(x + c)$ if the transform F satisfies either of the two following conditions:

(a) For every integer n, $F(n/p)$ is not zero.
(b) For some integer M, $F(n/p) \equiv 0$ for every $n > M$, and $F(n/p)$ is not zero for any $n \leq M$.

For example, condition (a) shows that the periodic image $f(x) = \sum_{n=-\infty}^{\infty} \Lambda\,[\pi(x - n)]$ is determined by its triple correlation [since the transform is $\sum_{n=-\infty}^{\infty} (1/\pi)\text{sinc}^2(n/\pi) \times \delta(u - n)$, which is positive at every n]. Condition (b) shows that the same is true of the periodic band-limited image obtained by convolving $\text{sinc}^2(x/w)$, $w > 1$, with the function f just defined.

Proof: If f has period p and $\ddot{a}_{2,g} = \ddot{a}_{2,f}$, then Eq. (4) shows that for all integers n, m at which $F(n/p)$, $F(m/p)$, and $F[(n + m)/p]$ are nonzero, we have

$$\begin{aligned} PhaG[(n + m)/p] - PhaF[(n + m)/p] \\ = PhaG(n/p) - PhaF(n/p) + PhaG(m/p) \\ - PhaF(m/p) + 2\pi N, \quad (31) \end{aligned}$$

where N is an arbitrary integer. Let

$$D(j) = PhaG(j/p) - PhaF(j/p)$$

for integers j at which $F(j/p)$ is nonvanishing. Then Eq. (31) implies that $D(n + m) = D(n) + D(m) + 2\pi N \times (n, m)$ for all $n, m, n + m$ where D is defined, with $N(n, m)$ an integer. If D is defined for all $j \leq M$, then, for $j \leq M$, $D(j) = jD(1) + 2\pi\sigma$, where σ is a sum of integers. Thus

$$\begin{aligned} PhaG(j/p) &= PhaF(j/p) + (j/p)[PhaG(1/p) - PhaF(1/p)] \\ &= PhaF(j/p) + 2\pi c(j/p) + 2\pi\sigma \end{aligned}$$

for $j \leq M$. So if $F(n/p) \neq 0$ for all n or for all $n \leq M$ with $F(u) \equiv 0$ for $u > M/p$, we have $G(u) = \exp(i2\pi cu)F(u)$ for all u, and $g(x) = f(x + c)$.

F. Triple Correlation Nonuniqueness and Image Reconstruction

Subsection 2.D.2 showed that, in principle, any image of finite size can be reconstructed from its triple-correlation function in two different ways: (i) determine the amplitude spectrum directly from the bispectrum by using Eq. (11) and the phase spectrum recursively by using Eq. (12) and (ii) determine the cumulants, and from them the moments of the image, from the derivatives of the bispectrum at zero (or, equivalently, from moments of the triple correlation), and reconstruct the image transform from the moments. Method (ii) cannot be applied to infinitely extended periodic images (which always have infinite moments) or to the integrable images (21) and (22) whose triple correlations were shown in Subsection 2.D.3 to be nonunique (since their even moments are all infinite). But there is no immediate reason why method (i) cannot be successfully applied to periodic images (for example, those that satisfy either condition of Theorem 7) or to some integrable images with infinite support even if they happen to have infinite moments (e.g., those covered by Theorem 3 or 4). However, it must fail somehow for images whose triple correlations do not have an inverse that is unique up to translation. This section explains the relationship between image reconstruction by method (i) and the infinite images whose triple correlations were shown in Subsections 2.D.3 and 2.E to be nonunique. We show first how the method fails for such images in cases in which the bispectrum of the entire image is assumed to be available and then discuss what happens when it is applied to the bispectrum of a finite portion of an image whose infinite version has a nonunique bispectrum. In the latter case Theorem 1 implies that the bispectrum of the finite portion always has an inverse that is unique up to translation, and in principle that inverse should be recoverable by method (i) no matter what the size of the observation window is. However, since this is not true in the limit, one expects recovery to become more difficult in some sense as the size of the window grows. We show by example the form that this difficulty takes.

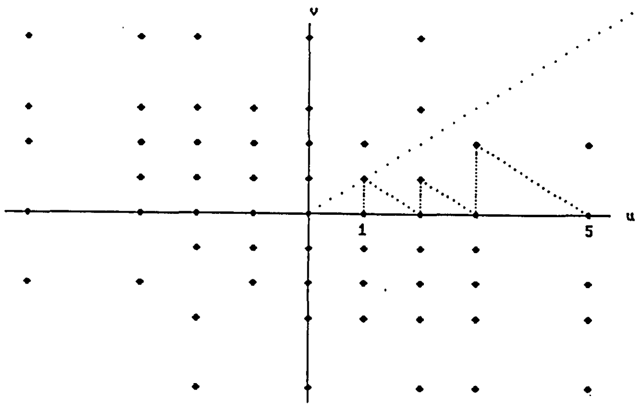

Fig. 3. Geometrical interpretation of recursive-phase reconstruction from the bispectrum. Heavy dots represent points in the u,v plane where the bispectrum of a hypothetical 1-D image is nonzero. The zigzag dotted line connecting dots in the region $0 \le u, 0 \le v \le u$ shows the recursive path for reconstructing the phase at frequency $u = 5$.

Figure 3 illustrates the geometry of recursive reconstruction of the phase spectrum of an image from its bispectrum. Large dots in the u,v plane represent points where the bispectrum of a hypothetical 1-D image f is nonzero. Recall that the symmetry relations (vi)–(viii) in Subsection 2.C imply that all the information in the bispectrum is contained in the octant $0 \le u, 0 \le v \le u$. We focus on that region, i.e., the area bounded by the u axis and the dotted line representing the $v = u$ diagonal. The amplitude spectrum $|F(u)|$ can be recovered from values of the bispectrum $A_{2,f}(u,v)$ along the u axis by means of Eq. (11). Here $|F(u)| > 0$ at $u = 0, 1, 2, 3, 5$. The problem, then, is to recover $PhaF(u)$ at those frequencies. The phase at $u = 1$ can be set arbitrarily, and $PhaF(u) - PhaF(1)$ for $u = 2, 3, 5$ can be reconstructed recursively by using Eq. (12), which corresponds to the zigzag dotted path connecting $(1, 0)$ to $(5, 0)$. For any given frequency u', a necessary condition for $PhaF(u')$ to be determined by the phase at some lower-frequency p by means of recursion (12), i.e., by

$$PhaF(u') = PhaF(p) + PhaF(u' - p) - PhaA_{2,f}(p, u' - p),$$

is that the bispectrum $A_{2,f}(u,v)$ be nonzero at the point $(p, u' - p)$, so that $PhaA_{2,f}(p, u' - p)$ is defined, and at $(p, 0)$ and $(u' - p, 0)$, so that $PhaF(p)$ and $PhaF(u' - p)$ are defined. [$PhaF(u)$ is defined only if $A_{2,f}(u, 0) \neq 0$, since $A_{2,f}(u, 0) = F^2(u)F(0)$.] If there is no frequency p with $u'/2 \le p < u'$ for which all three conditions are satisfied, then the phase at u' cannot be recursively reconstructed from any lower frequencies. For a given u' one can quickly check whether there is any p that satisfies the first two conditions by examining the bispectrum along the line from $(u', 0)$ to the $v = u$ diagonal and at the point on the u axis directly below any point $(p, u' - p)$ at which $A_{2,f}(p, u' - p) \neq 0$.

Figures 4 and 5 illustrate how recursive-phase reconstruction fails for images whose triple correlations do not have inverses that are unique up to translation. The upper part of Fig. 4 shows the support (for $u,v \ge 0$) of the bispectrum of the infinitely extended periodic image $f(x) = 2 + \cos 2\pi x + \cos 6\pi x$, whose triple correlation was shown in Subsection 2.E to be the same as that of $2 + \cos 2\pi x - \cos 6\pi x$. [The lower graph in Fig. 4 shows the normalized amplitude spectrum of the image, i.e., $|F(u)|/F(0)$.] It can be seen that the frequency $u = 3$ fails the test just described: $PhaF(3)$ cannot be reconstructed from $PhaF(1)$ because the line from $(3, 0)$ to the $v = u$ diagonal is empty: $A_{2,f}(p, 3 - p)$ is zero for all p. Thus the bispectrum can tell us nothing about $PhaF(3) - PhaF(1)$.

Figure 5 (heavily striped areas) shows the support of the bispectrum of the integrable image $\text{sinc}^2(x)(1 + \cos 6\pi x)$, whose triple correlation is the same as that of $\text{sinc}^2(x)(1 + \sin 6\pi x)$. Here the amplitude spectrum $|F(u)|$ (shown in the bottom part of the figure) is nonzero on the intervals $(0, 1)$ and $(2, 4)$, but there is no frequency u' in the interval $(2, 4)$ whose phase can be reconstructed from that of any frequency $p < 2$, because, if $u' > 2$, either $A_{2,f}(p, u' - p)$ is zero or $p > 2$. Thus in this case the frequencies in the ranges $(0, 1)$ and $(2, 4)$ form isolated islands: Knowing only the bispectrum and the phases of frequencies in one island, one can infer nothing about the phase of any frequency in the other island.

Figures 6 and 7 illustrate what happens when one attempts to recover from the bispectrum the phase spectrum of a finite portion of an infinite image whose triple correlation does not have a unique inverse. Here the infinite image is the periodic function $f(x) = 2 + \cos 2\pi x + \cos 6\pi x$. Suppose first that we observe $f(x)$ through a window of width 1.0 and treat the image as zero outside that window, so that we compute the triple correlation of

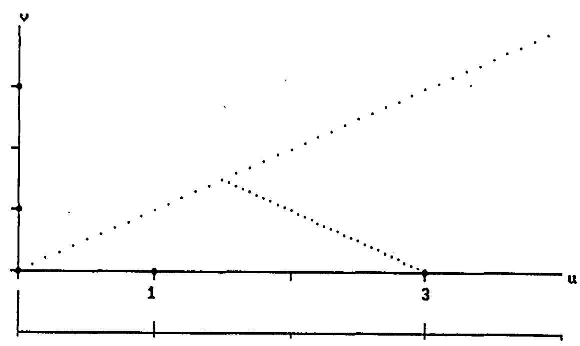

Fig. 4. Geometrical reason why the phase spectrum of the infinite periodic image $2 + \cos 2\pi x + \cos 6\pi x$ cannot be reconstructed from its bispectrum. Top: bispectrum for $u,v \ge 0$; heavy dots indicate points of nonzero amplitude. Bottom: normalized amplitude spectrum of the image.

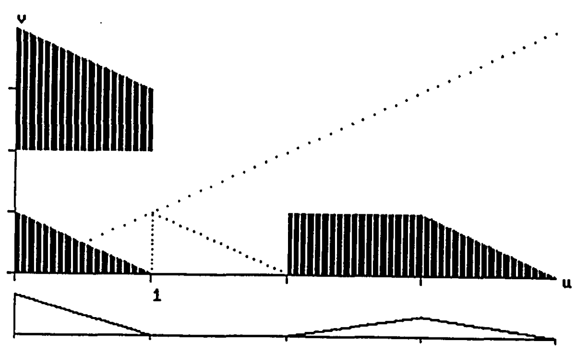

Fig. 5. Geometrical reason why the phase spectrum of the infinite integrable image $\text{sinc}^2(x)(1 + \cos 6\pi x)$ cannot be reconstructed from its bispectrum. Top: heavy striped areas are regions of the u,v plane ($u,v \ge 0$) where the bispectrum is nonzero. Bottom: normalized amplitude spectrum of the image.

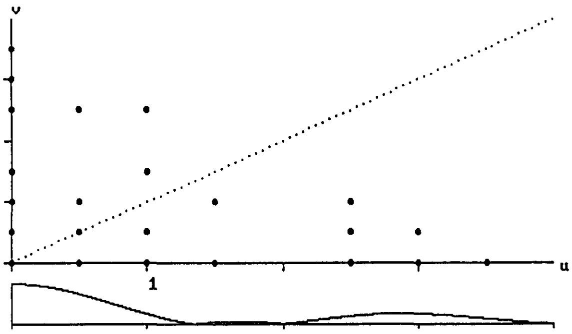

Fig. 6. Top: bispectrum of a windowed version of the image $2 + \cos 2\pi x + \cos 6\pi x$ for a narrow window (width = 1). Heavy dots indicate points where the normalized bispectrum of the windowed image has an amplitude of 0.001 or more. Sampling interval, 0.50; cutoff, 0.001. Bottom: normalized amplitude spectrum of the windowed image.

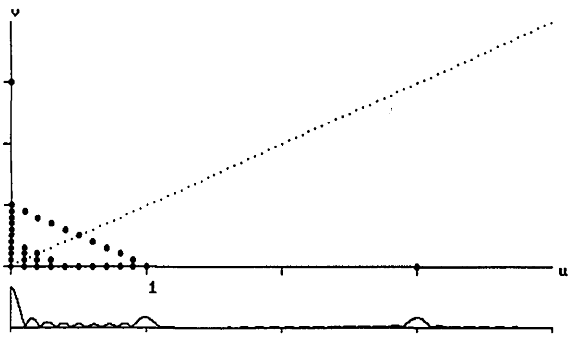

Fig. 7. Top: bispectrum of the windowed version of $2 + \cos 2\pi x + \cos 6\pi x$ for a window of width 9. Sampling interval, 0.10; cutoff, 0.001. Bottom: normalized amplitude spectrum of the windowed image.

the finite image $fw(x) = \text{rect}(x)f(x)$ [using definition (3) in Section 1]. The upper graph in Fig. 6 shows the support of the bispectrum of fw: The large dots indicate the points where the normalized bispectrum $A_{2,fw}(u,v)/A_{2,fw}(0,0)$ has an absolute value of ≥ 0.001. (The bispectrum has been sampled at twice the Nyquist rate. The lower graph shows the normalized amplitude spectrum of fw.) Examination of the u axis shows that the normalized power spectrum of fw [i.e., $A_{2,fw}(u,0) = |FW(u)|^2/FW(0)$] is ≥ 0.001 at frequencies 0.5, 1.0, 1.5, 2.5, 3.0, and 3.5; and one can see from the arrangement of dots in the u,v plane that, when $PhaFW(0.5)$ is fixed arbitrarily, the phases at the other frequencies can be recursively determined from the bispectrum. Thus both the amplitudes and the phases of the significant frequency components of $fw(x)$ can be reconstructed from its bispectrum.

However, as the window widens the situation deteriorates. Figure 7 shows the support of the bispectrum of $fw(x) = \text{rect}(x/9)f(x)$, i.e., $f(x)$ seen through a window of width 9. Again, the large dots indicate the points where the bispectrum of fw is ≥ 0.001. (The sampling rate has been increased to 10, still greater than the Nyquist rate.) We see from the u axis that the normalized power spectrum exceeds 0.001 at $u = 3.0$, 1.0, and numerous points below 1.0, but the bispectrum contains no points with absolute values ≥ 0.001 along the line from (3,0) to the $v = u$ diagonal. Consequently, $PhaFW(3)$ cannot be determined recursively from any lower frequencies, that is, from any lower frequencies whose normalized spectral powers are at least 0.001. In particular, there is no recursive path from $u = 1$ to $u = 3$ connecting the two frequencies that have sizable spectral power. Of course, since fw has bounded support, we know that in principle $PhaFW(3)$ must be recursively determinable by some sequence of frequencies beginning at $u = 1$. However, it must be a path along which the absolute values of the bispectrum are all near zero. In a real-world problem such values might be zero in fact, spuriously inflated by noise or round-off error. The problem is that, if we use the phases of such suspect points to reconstruct the phase recursively at a frequency that does have significant spectral power, such as $u = 3$ in this example, their small absolute values are irrelevant: the recursion treats the phases of all frequencies with equal respect, regardless of their spectral amplitudes. But the reliability of the phase that it assigns to any frequency depends on the reliability of the phases assigned to earlier frequencies. If all the possible recursive paths from one significant frequency to another involve intermediate frequencies whose spectral amplitudes are very small and thus likely to be actually zero except for measurement error, the reliability of subsequent phase assignments must suffer accordingly.

From a practical standpoint, then, the fundamental problem posed by the nonuniqueness of the triple correlations of infinite images is essentially statistical. In principle, with perfect measurement and computation, any finite portion of any infinite image can be unambiguously reconstructed from its bispectrum, even if the full image does not have a unique inverse. But in the latter case it is precisely the gaps in the bispectrum of the full image that make its triple correlation noninvertable, and the bispectrum of a windowed version of such an image must converge to zero in those gap regions as the size of the window increases. This forces the recursion linking frequencies across a gap to depend on intermediate frequencies whose true spectral power is becoming vanishingly small and, in the presence of noise, increasingly difficult to distinguish reliably from zero. Consequently, phase reconstruction for such an image becomes increasingly unreliable as the observation window widens: in effect, the more we see of its bispectrum, the less confidence we have in our reconstruction of the image.

3. HIGHER-ORDER AUTOCORRELATIONS OF 2-D IMAGES

A. Definitions and Basic Properties

We call a function $f:\mathbf{R}^2 \rightarrow \mathbf{R}$ an image function (or, simply, an image) if f is nonnegative and integrable over every finite rectangle.[24] We write (x_1, x_2) as X and denote the integral of $f(X)$ over the rectangle R by $\int_R f(X)\mathrm{d}X$. If $\int_R f(X)\mathrm{d}X = \int_R g(X)\mathrm{d}X$ for every R, we say $f = g$. If $\int_{\mathbf{R}^2} f(X)\mathrm{d}X = 0$, then $f = 0$. An image f is integrable if $\int_{\mathbf{R}^2} f(X)\mathrm{d}X$ is finite. In that case the kth-order ACF of f is defined to be

$$a_{k,f}(S_1,\ldots,S_k) = \int_{\mathbf{R}^2} f(X)f(X + S_1)\ldots f(X + S_k)\mathrm{d}X, \tag{32}$$

where $S_j = (s_{j,1}, s_{j,2})$ for $j = 1,\ldots,k$. [For $k = 2$, Eq. (32) defines the triple correlation of an integrable 2-D image.] A probabilistic argument like the one given in Subsection 2.A shows that, for every k, $a_{k,f}$ is integrable over $\mathbf{R}^{2k}$. The Fourier transform F is

$$F(U) = \int_{\mathbf{R}^2} \exp(-i2\pi U \cdot X) f(X) \mathrm{d}X,$$

where $U = (u_1, u_2)$ and $U \cdot X = u_1 x_1 + u_2 x_2$. The transform of $a_{k,f}$ is again denoted $A_{k,f}$. ($A_{2,f}$ is the 2-D bispectrum.) A calculation like the one leading to Eq. (5) shows that

$$A_{k,f}(U_1, \ldots, U_k) = F\left(-\sum_{j=1}^{k} U_j\right) \prod_{j=1}^{k} F(U_j), \tag{33}$$

where $U_j = (u_{j,1}, u_{j,2})$, $j = 1, \ldots, k$. It follows from Eq. (33) that two integrable 2-D images f and g have the same kth-order ACF if and only if their transforms satisfy

$$F\left(-\sum_{j=1}^{k} U_j\right) \prod_{j=1}^{k} F(U_j) = G\left(-\sum_{j=1}^{k} U_j\right) \prod_{j=1}^{k} G(U_j) \tag{34}$$

for all arguments.

Periodicity is a simple concept on the line but not in the plane. We follow Klein and Tyler[4] and confine our analysis to periodic images constructed by defining an arbitrary image function within a square region centered at the origin and then tiling the plane with it. Let ω_p be the open square $\{(x_1, x_2): |x_1| < p/2, |x_2| < p/2\}$, and let f_p be an image function defined on ω_p, with $f_p(X) \equiv 0$ for $X \notin \omega_p$. Then f is a periodic image if, for some ω_p and f_p,

$$f(X) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f_p[X + (np, mp)].$$

The kth-order ACF $\ddot{a}_{k,f}$ of a periodic image f is defined to be

$$\ddot{a}_{k,f}(S_1, \ldots, S_k) = (1/p^2) \int_{\omega_p} f(X) f(X + S_1) \ldots f(X + S_k) \mathrm{d}X. \tag{35}$$

The Fourier transform $\ddot{A}_{k,f}$ of $\ddot{a}_{k,f}$ is related to the transform of f in the same way as in the 1-D case:

$$\ddot{\mathbf{A}}_{k,f}(U_1, \ldots, U_k) = (1/p^2) F_p\left(-\sum_{j=1}^{k} U_j\right) \prod_{j=1}^{k} F_p(U_j) III(pU_j),$$

where $F_p(U)$ is the transform of a single period of f,

$$F_p(U) = \int_{\omega_p} \exp(-i2\pi U \cdot X) f_p(X) \mathrm{d}X,$$

and III denotes the 2-D Dirac comb,

$$III(U) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \delta(u_1 - n, u_2 - m).$$

$\ddot{A}_{k,f}$ is concentrated at the $2k$-dimensional lattice points $(U_1, \ldots, U_k) = [(n_{1,1}/p, n_{1,2}/p), \ldots, (n_{k,1}/p, n_{k,2}/p)]$, where the $n_{j,i}$ are all integers, and at such points the value of its delta is

$$(1/p^2)^{k+1} F_p\left(-\sum_{j=1}^{k} U_j\right) \prod_{j=1}^{k} F_p(U_j).$$

The transform of f is $F(U) = (1/p^2) F_p(U) \Sigma_n \Sigma_m \delta(u_1 - n/p, u_2 - m/p)$, so the following 2-D analog to Eq. (6) is valid:

$$\ddot{\mathbf{A}}_{\mathbf{k},\mathbf{f}}(U_1, \ldots, U_k) = F\left(-\sum_{j=1}^{k} U_j\right) \prod_{j=1}^{k} F(U_j). \tag{36}$$

It follows from Eq. (36) that two periodic 2-D images f and g have the same kth-order ACF if and only if their transforms satisfy Eq. (34).

Using Eqs. (33) and (36) in the same way that Eqs. (5) and (6) were used in Subsection 2.C, one can quickly show that ACF's have the same basic properties in two dimensions as in one dimension: If f and g are integrable or periodic images, then

(i′) If $a_{k,g} = a_{k,f}$ (or $\ddot{a}_{k,g} = \ddot{a}_{k,f}$), then $a_{j,g} = a_{j,f}$ ($\ddot{a}_{j,g} = \ddot{a}_{j,f}$) for all $1 \le j < k$. (The kth-order ACF determines all the lower orders.)

(ii′) If $g(X) = f(X + C)$, where $C = (c_1, c_2)$, then $a_{k,g} = a_{k,f}$ ($\ddot{a}_{k,g} = \ddot{a}_{k,f}$) for all k. (Translating an image leaves all its ACF's unchanged.)

(iii′) If $a_{k,g} = a_{k,f}$ ($\ddot{a}_{k,g} = \ddot{a}_{k,f}$) and h is any third image, the convolutions $h * g$ and $h * f$ have identical kth-order ACF's.

(iv′) If images f and g have the same kth-order ACF, so do $f * f$ and $g * g$.

(v′) If images $f(x, y)$ and $g(x, y)$ have the same kth-order ACF, so do $f(ax, by)$ and $g(ax, by)$ for any constants a and b.

(A prime attached to a statement number here indicates that it is a 2-D version of an earlier 1-D result with the same number.)

For the triple correlation, the following symmetry relations, analogous to relations (vi)–(viii) of Subsection 2.C, follow from Eqs. (33) and (36):

(vi′)
$$A_{2,f}(U, V) = A_{2,f}(V, U),$$
$$\ddot{A}_{2,f}(U, V) = \ddot{A}_{2,f}(V, U);$$

(vii′)
$$A_{2,f}(U, V) = A_{2,f}(U, -U - V),$$
$$\ddot{A}_{2,f}(U, V) = \ddot{A}_{2,f}(U, -U - V);$$

(viii′)
$$A_{2,f}(U, V) = A_{2,f}^*(-U, -V),$$
$$\ddot{A}_{2,f}(U, V) = \ddot{A}_{2,f}^*(-U, -V).$$

B. Uniqueness for Integrable 2-D Images

1. *Finite Images*

For 2-D images with bounded support in the plane we have the following analog of Theorem 1:

Theorem 1′. Suppose that f is an integrable image and that, for some real number $b > 0$, $f(X) \equiv 0$ for $|X| \ge b$. Then $a_{2,g} = a_{2,f}$ for another image g if and only if $g(X) = f(X + C)$ for some constant $C = (c_1, c_2)$.

Proof: "If" follows from property (ii′) of Subsection 3.A above. The proof of "only if" follows the same lines as in Theorem 1. Here $a_{2,g} = a_{2,f}$ implies [from Eq. (33)] that the bispectra of f and g satisfy the relationship

$$G(U)G(V)G(-U - V) = F(U)F(V)F(-U - V) \tag{37}$$

for all U, V. Thus $F(0) = G(0)$ and, in general, $|F(U)| = |G(U)|$. If $F(0) = \int_{\mathbf{R}^2} f(X) \mathrm{d}X = 0$, then $f = 0$, and g must be 0 also. We assume, then, that $F(0) \ne 0$, and without loss of generality we assume that $F(0) = 1 = G(0)$.

Then f and g are probability densities of random vectors $\mathbf{X}_f = (\mathbf{x}_{f,1}, x_{f,2})$, $\mathbf{X}_g = (\mathbf{x}_{g,1}, \mathbf{x}_{g,2})$, and F and G are the characteristic functions of these densities. We rely again on two general properties of characteristic functions: Every characteristic function is continuous and equals 1.0 at the origin (so F and G are both nonvanishing in some neighborhood of the origin); and, if F is the characteristic function of a density with bounded support in the plane, F is completely determined by its values in a neighborhood of the origin. {For every fixed value of U, $U \cdot \mathbf{X}_f$ is a 1-D random variable whose density has bounded support on the line, so the characteristic function $\phi(t) = E[\exp(-i2\pi tU \cdot \mathbf{X}_f)] = F(tU)$ is determined for all t by its derivatives at $t = 0$, i.e., by the values of $F(U)$ in a neighborhood of the origin. Thus $\phi(1) = F(\mathrm{U})$ is so determined, and this is true for all U.} Consequently, if $G(U) = \exp(i2\pi C \cdot U)F(U)$ in a neighborhood of the origin for some constant C, then the density $g(X)$ must equal $f(X + C)$. Since $F(U)$ and $G(U)$ are nonvanishing in some neighborhood $|U| < \beta$, Eq. (37) implies that, for all $|U|, |V| < \beta/2$,

$$PhaG(U) + PhaG(V) - PhaG(U + V) = PhaF(U) + PhaF(V) - PhaF(U + V) + 2\pi N(U, V), \tag{38}$$

with $PhaG(U)$ and $PhaF(U)$ both continuous and equal to 0 at $U = 0$. It follows that $N(U, V) = 0$ for $|U|, |V| < \beta/2$. Writing $D(U) = PhaG(U) - PhaF(U)$ and rearranging Eq. (38), we have the 2-D Cauchy functional equation

$$D(U + V) = D(U) + D(V), \tag{39}$$

which holds for all U, V in a neighborhood of the origin, with D continuous. Aczel[22] shows that all continuous solutions to Eq. (39) take the form $D(U) = B \cdot U$, where $B = (b_1, b_2)$ is a constant, so for some constant C, $PhaG(U) = PhaF(U) + 2\pi C \cdot U$. Thus, in a neighborhood of the origin, $G(U) = \exp(i2\pi C \cdot U)F(U)$, so for all X, $g(X) = f(X + C)$.

2. Infinite Images

The 1-D counterexamples (16) and (17) in Subsection 2.D.3 also show that for every k there are 2-D images with infinite support that have identical kth-order ACF's but that are not translations of each other. Consequently, the only completely general uniqueness theorem possible here is the following:

Theorem 2′. If f and g are integrable images and $a_{k,g} = a_{k,f}$ for all k, then $g(X) = f(X + C)$ for some constant C.

Proof: When the vectors U and V are substituted for arguments u and v, the proof is the same as that of Theorem 2.

The following triple-correlation-uniqueness theorems for special classes of infinite images are 2-D versions of Theorems 3–5:

Theorem 3′. If f is an integrable image and $a_{2,g} = a_{2,f}$ for another image g, then $g(X) = f(X + C)$ if the Fourier transform of f satisfies any of the following conditions:

(a) $F(U)$ is nonvanishing for all U.

(b) For every line L_θ through the origin of the spatial frequency plane [$L_\theta = \{(r \cos\theta, r \sin\theta): -\infty < r < \infty\}$ for a fixed $\theta \in (0, \pi)$], there is a frequency cutoff $b_\theta \ge 0$ such that $F(r \cos\theta, r \sin\theta) \equiv 0$ for $|r| \ge b_\theta$, and $F(r \cos\theta, r \sin\theta) \ne 0$ for $r < b_\theta$. (That is, f is band limited, and for every spatial-frequency orientation its spectrum is nonvanishing below the cutoff for that orientation.)

(c) For every line L_θ through the origin of the spatial-frequency plane there is a cutoff b_θ such that $F(r \cos\theta, r \sin\theta) \equiv 0$ for $|r| \ge b_\theta$, and $F(r \cos\theta, r \sin\theta) = 0$ for at most a finite number of values of $r < b_\theta$.

(d) For every line L_θ through the origin of the spatial-frequency plane, $F(U)$ has at most a finite number of zeros in every finite interval of L_θ.

Condition (a) covers, for example, 2-D Gaussians and Gabor functions and the 2-D Cauchy function $f(x_1, x_2) = (1 + x_1^2 + x_2^2)^{-3/2}$. Condition (b) shows that $\mathrm{sinc}^2(x)\,\mathrm{sinc}^2(y)$ and $J_1^2(r)/r^2$, $r^2 = x_1^2 + x_2^2$, are determined up to translation by their triple correlations. These are, respectively, the incoherent impulse responses of diffraction-limited optical systems with square and circular exit pupils.[25] Condition (c) shows that the images of finite objects formed by such systems are determined up to translation by their triple correlations. This is so because, along any line through the origin, the transform $F(U)$ of a 2-D image of finite size has at most a finite number of zeros. [Suppose that $F(0) \ne 0$ and f has bounded support in the plane. For any fixed θ, $\Phi(r) = F(r \cos\theta, r \sin\theta)/F(0)$ is the characteristic function of a 1-D random variable $(\cos\theta)\mathbf{x}_1 + (\sin\theta)\mathbf{x}_2$ whose density has bounded support on the line, so $\Phi(r)$ has at most finitely many zeros in any finite interval.] Condition (d) shows that any finite image blurred by a Gaussian (Gabor, Cauchy) impulse response is determined by its triple correlation.

Proof: The proof of Theorem 1′ shows that there is a neighborhood of the origin in which $G(U) = \exp(i2\pi C \cdot U)F(U)$ for some constant $C = (c_1, c_2)$. For each fixed value of θ in $[0, \pi)$ let $U_\theta = (\cos\theta, \sin\theta)$, and write $U = rU_\theta$ for U on the line L_θ. Then $\Phi(r) = F(rU_\theta)$ and $\Gamma(r) = G(rU_\theta)$ are 1-D Fourier transforms and, for all p, q,

$$\Gamma(p)\Gamma(q)\Gamma(-p - q) = \Phi(p)\Phi(q)\Phi(-p - q).$$

Consequently, the Fourier inverses ϕ and γ of Φ and Γ are 1-D images with $a_{2,\gamma} = a_{2,\phi}$. If F satisfies any one of conditions (a)–(d) here, Φ satisfies the same condition in Theorem 3, and thus for some real constant c_θ, $\Gamma(r) = \exp(i2\pi c_\theta r)\Phi(r)$ for all r. So $G(rU_\theta) = \exp(i2\pi c_\theta r)F(rU_\theta)$ for all r. But in a neighborhood of zero we also have $G(rU_\theta) = \exp(i2\pi C \cdot rU_\theta)F(rU_\theta)$, so $rc_\theta = C \cdot rU_\theta$, and for all U on the line L_θ, $G(U) = \exp(i2\pi C \cdot U)F(U)$. This is true for every θ, so $g(X) = f(X + C)$.

The next theorem involves a 2-D version of the finite linking condition defined above in connection with Theorem 4. As before, Δ_f is the neighborhood of the origin (now guaranteed by the proof of Theorem 1′), where F and G are nonvanishing and $G(U) = \exp(i2\pi C \cdot U)F(U)$. Here if $F(U) \ne 0$, with U lying on a line L_θ through the origin, U is said to be finitely linked to Δ_f if there is a sequence of points on L_θ, $P_0 \in \Delta_f$, $P_0 < P_1 < P_2 < \ldots < P_{n-1} < P_n = U$, with $P_i = (1 + \alpha_1 + \cdots + \alpha_i)P_0$, $0 < \alpha_i \le 1$, and $F(P_i) \ne 0$ for $i = 1, \ldots, n$. Then the proof of Theorem 4 transfers immediately to two dimensions, and we have the following:

Theorem 4′. If f is an image for which each point U, where $F(U) \neq 0$, can be finitely linked to the neighborhood Δ_f, and if $a_{2,g} = a_{2,f}$ for another image g, then $g(X) = f(X + C)$ for some constant C.

Finally, we prove a 2-D version of Theorem 5. The moment sequence ($\mu_{n,m}$: $n = 0,1,\ldots$; $m = 0,1,\ldots$) of an image $f(X)$ is defined by $\mu_{n,m} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^n x_2^m f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2$, and f is said to be determined by its moments if it is the only image function that has the moment sequence ($\mu_{n,m}$). If $F(u_1, u_2)$ is the Fourier transform of f and $F^{(n,m)}(u_1,u_2) = \partial^n\partial^m/\partial u^n \partial v^m F(u_1,u_2)$, then, whenever it exists, $\mu_{n,m} = F^{(n,m)}(0,0)(i2\pi)^{-n-m}$. Let $\Phi(u,v) = \log F(-u_1/2\pi, -u_2/2\pi)$. The cumulants $\kappa_{n,m}$ of f are the quantities $(i^{-n-m})\Phi^{(n,m)}(0,0)$, and, just as in the 1-D case, the cumulants of f uniquely determine its moments and vice versa. Thus f is determined by its moments if and only if it is determined by its cumulants. And if $f(X)$ is determined by its cumulants, so too is $f(X + C)$ for any constant $C = (c_1, c_2)$, since the cumulants of the latter are the same as those of $f(X)$ except for $\kappa_{1,0}$, which becomes $\kappa_{1,0} - c_1$, and $\kappa_{0,1}$, which becomes $\kappa_{0,1} - c_2$. So if $f(X)$ is determined by its moments, $f(X + C)$ is also. If f and g are images that have the same triple correlation, the proof of Theorem 1′ shows that, for some C, $G(U) = \exp(i2\pi C \cdot U)F(U)$ in a neighborhood of the origin, so $g(X)$ has the same moments as $f(X + C)$. Consequently, if f is an image that is uniquely determined by its moments, then f is uniquely determined up to translation by its triple correlation. We state this as Theorem 5′ below. Shohat and Tamarkin[17] discuss the general problem of characterizing the nonnegative real functions on $\mathbf{R}^2$ that are uniquely determined by their moments. We mention only one sufficient condition, which is analogous to the earlier condition Eq. (25). For that purpose, we define a sequence $\{M_n: n = 1,2,\ldots\}$ by $M_n = \int_{\mathbf{R}^2} |X|^n f(X)\mathrm{d}X$, where $|X| = (x_1^2 + x_2^2)^{1/2}$, and show that f is determined by its moments if

$$\lim_{n\to\infty} \sup|M_n/n!|^{1/n} = \lambda < \infty. \tag{40}$$

Theorem 5′. If f is an integrable image that is uniquely determined by its moments [for example, if f satisfies condition (40)] and g is another image with $a_{2,g} = a_{2,f}$, then $g(X) = f(X + C)$ for some constant C.

Proof: In view of the preceding discussion we need show only that Eq. (40) guarantees that f is determined by its moments. We know that, for some constant C, $G(U) = \exp(i2\pi C \cdot U)F(U)$ in a neighborhood of the origin, and as usual we assume without loss of generality that $F(0) = 1$. Then F is the characteristic function of a random vector $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2)$, and for every fixed value of $U = (u_1, u_2)$ and any real number t, $\Phi_U(t) = F(tU)$ is the characteristic function of the 1-D random variable $U \cdot \mathbf{X} = u_1\mathbf{x}_1 + u_2\mathbf{x}_2$. The moments ν_n of this random variable are determined by the sequence $\{\mu_{n,m}\}$, since

$$\nu_n = \int_{\mathbf{R}^2} (u_1x_1 + u_2x_2)^n f(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2$$

$$= \sum_{j=0}^{n} n!/[j!(n-j)!]u_1^{n-j}u_2^{j}\mu_{n-j,j}.$$

As in the proof of Theorem 5, the function $\exp(i2\pi tC \cdot U)\Phi_U(t)$ is uniquely determined by analytic continuation for all real t (in particular, for $t = 1$) by its values in a neighborhood of $t = 0$ (i.e., by its derivatives at 0) if the moments of $U \cdot \mathbf{X}$ are finite and satisfy the condition

$$\lim_{n\to\infty} \sup|\nu_n/n!|^{1/n} = \rho < \infty.$$

It remains to be shown that Eq. (40) implies that this condition is satisfied. Starting with

$$|\nu_n| = \left|\int_{\mathbf{R}^2} (U \cdot X)^n f(X)\mathrm{d}X\right| \leq \int_{\mathbf{R}^2} |U \cdot X|^n f(X)\mathrm{d}X$$

and writing $U = p(\cos\alpha, \sin\alpha)$ and $X = q(\cos\beta, \sin\beta)$, we have $U \cdot X = pq(\cos\alpha - \beta)$ and $|U \cdot X|^n \leq (pq)^n = |U|^n|X|^n$. Thus $|\nu_n/n!|^{1/n} \leq |U|\,|M_n/n!|^{1/n}$, so, from the assumption of the theorem,

$$\lim_{n\to\infty} \sup|\nu_n/n!|^{1/n} \leq |U|\lambda = \rho < \infty.$$

Consequently, $\exp(i2\pi C \cdot U)F(U)$ is uniquely determined by its values in a neighborhood of the origin, and $g(X) = f(X + C)$.

The cumulants method described in Subsection 2.D for obtaining the moments of a 1-D image f from the derivatives of $A_{2,f}(-u/2\pi, -u/2\pi)$ at $u = 0$ can be extended in a straightforward way to 2-D images, using the derivatives at $t = 0$ of $A_{2,f}(tU, tU)$. Thus, in the same way that the 2-D bispectrum $A_{2,f}(u,v)$ of any 1-D image determined by its moments is completely characterized by its values along the line $v = u$ in a neighborhood of $u = 0$, the four-dimensional bispectrum $A_{2,f}(U,V)$ of any 2-D image determined by its moments is completely characterized by its values in the plane $V = U$ in a neighborhood of $U = (0,0)$.

It should be apparent that the arguments used to prove Theorems 1′–5′ do not depend on X being in $\mathbf{R}^2$. If we began instead with nonnegative real functions on $\mathbf{R}^n$ and rephrased the conditions of the theorems in the obvious ways, their proofs would still hold. Characteristic functions of any dimension are always continuous and nonvanishing in a neighborhood of the origin; the Cauchy functional equation [Eq. (39)] for $U, V \in \mathbf{R}^n$ always has the solution $D(U) = B \cdot U$ with B a constant in $\mathbf{R}^n$, and probability-density functions with bounded support are always uniquely determined by the values of their characteristic functions in a neighborhood of the origin; so Theorem 1′ is valid for $f: \mathbf{R}^n \to \mathbf{R}$. Theorem 2′ clearly does not depend on dimensionality, and the proofs of Theorems 3′–5′ rely on reducing 2-D statements to 1-D ones in ways that will work as well for n dimensions. Thus Theorem 1′, for example, shows that nonnegative integrable functions of time and space with bounded support in $\mathbf{R}^3$ are determined up to an arbitrary translation by their triple correlations, and condition (c) of Theorem 3′ shows that this remains true of such functions after they are subjected to low-pass filtering.

C. Uniqueness for Periodic 2-D Images

Examples (26) and (27) in Subsection 2.E show that for every k there are 2-D periodic images that are not translations of each other but that have identical kth-order ACF's. Consequently, the only completely general uniqueness result that one can hope for here is that every periodic image is determined up to translation by its entire set of ACF's of all orders. As in the 1-D periodic case, we do not know whether that is true,[19] so we prove instead a weaker result analogous to the earlier Theorem 6:

Theorem 6′. Suppose that f is a periodic 2-D image with period p and $F(1/p, 0)$ and $F(0, 1/p)$ are both nonzero. Then $\ddot{a}_{k,g} = \ddot{a}_{k,f}$ for all k for another image g if and only if $g(X) = f(X + C)$ for some constant C.

Proof: "If" is property (ii′) in Subsection 3.A. To show "only if," we take $p = 1$ for convenience. From $\ddot{a}_{2,g} = \ddot{a}_{2,f}$ and Eq. (34) we have $|F(U)| = |G(U)|$ for all U. Since $F(1, 0) \neq 0$, neither $F(-1, 0)$ nor $G(-1, 0)$ is zero, and we can write

$$G(-1,0) = |F(-1,0)|\exp\{i[PhaG(-1,0) - PhaF(-1,0) + PhaF(-1,0)]\}$$
$$= F(-1,0)\exp(-i2\pi c_1),$$

where $c_1 = PhaF(-1,0) - PhaG(-1,0)$. So

$$F(-1,0)/G(-1,0) = \exp(i2\pi c_1). \tag{41}$$

Similarly,

$$F(0,-1)/G(0,-1) = \exp(i2\pi c_2), \tag{42}$$

where $c_2 = PhaF(0,-1) - PhaG(0,-1)$. Then, since $\ddot{A}_{n+m,g} = \ddot{A}_{n+m,f}$ for all integers n, m, Eq. (34) with arguments $U_1 = \ldots = U_n = (-1,0)$, $U_{n+1} = \ldots = U_{n+m} = (0,-1)$, implies that

$$[G(-1,0)]^n[G(0,-1)]^m G(n,m) = [F(-1,0)]^n[F(0,-1)]^m F(n,m),$$

and division using Eqs. (41) and (42) yields

$$G(n,m) = \exp[i2\pi(c_1 n + c_2 m)]F(n,m).$$

Thus $G(U) = \exp(i2\pi C \cdot U)F(U)$ for all U, and $g(X) = f(X + C)$.

We conclude with a 2-D version of Theorem 7:

Theorem 7′. If f is a periodic image with period p and $\ddot{a}_{2,g} = \ddot{a}_{2,f}$ for another image g, then $g(X) = f(X + C)$ if f satisfies either one of the two following conditions:

(a) $F(n/p, m/p) \neq 0$ for all integers n and m.
(b) For some integers N and M, $F(n/p, m/p) \neq 0$ if $|n| \leq N$ and $|m| \leq M$, and $F(n/p, m/p) \equiv 0$ otherwise.

Proof: Let $p = 1$, and suppose that $\ddot{A}_{2,f}$ is given. Then $|F(n, m)|$ is determined by $\ddot{A}_{2,f}[(n,m),(-n,-m)]$. The phases $PhaF(n, 0)$, $PhaF(0, m)$ can be constructed recursively from $Pha\ddot{A}_{2,f}$ up to arbitrary values for $PhaF(1,0)$ and $PhaF(0,1)$ by using the relations

$$PhaF(j+1,0) = Pha\ddot{A}_{2,f}[(-j,0),(-1,0)] + PhaF(j,0) + PhaF(1,0),$$

$$PhaF(0,k+1) = Pha\ddot{A}_{2,f}[(0,-k),(0,-1)] + PhaF(0,k) + PhaF(0,1).$$

Finally,

$$PhaF(n,m) = Pha\ddot{A}_{2,f}[(-n,0),(0,-m)] + PhaF(n,0) + PhaF(0,m),$$

so $F(n, m)$ is determined in terms of $F(1, 0)$, $F(0, 1)$ for all n, m for condition (a) or for $n \leq N, m \leq M$ for condition (b).

APPENDIX A: PROOF OF THEOREM 3

For condition (a) of Theorem 3, suppose that f and g have the same triple correlation and that the transform F never vanishes. Then from Eq. (8), $|G(u)| = |F(u)|$ for all u, so G is also nonvanishing, and the phase relationship [Eq. (9)] holds for all u and v. The continuity argument given in the proof of Theorem 1 then shows that $N(u,v) = 0$ for all u,v. Consequently, functional Eq. (10) is valid for all u,v and $PhaG(u) = PhaF(u) + 2\pi cu$ for all u, so $G(u) = \exp(i2\pi cu)F(u)$ is valid for all u and $g(x) = f(x + c)$.

To prove condition (b), suppose that f is a band-limited image whose transform is nonvanishing within its bandwidth; i.e., $F(u) \equiv 0$ for $|u| \geq b$, and $F(u) \neq 0$ for $|u| < b$. If g is another image with the same triple correlation as f, then $G(u)$ is also nonvanishing for $|u| < b$ and zero for $|u| \geq b$, and Eq. (9) is valid for all $-b/2 < u,v < b/2$, with $N(u,v) \equiv 0$. Consequently, Eq. (10) holds for u,v in $(-b/2, b/2)$, so $PhaG(u) = PhaF(u) + 2\pi cu$ in that interval, and by continuity this is true also at $u = \pm b/2$. Then for $b/2 \leq u < b$ we write $u = u/2 + u/2$, and Eq. (10) implies that $PhaG(u) = PhaF(u) + 2\pi c(u/2) + 2\pi c(u/2)$. So $PhaG(u) = PhaF(u) + 2\pi cu$ for all u in $(-b, b)$, and thus $G(u) = \exp(i2\pi cu)F(u)$ for all u, so $g(x) = f(x + c)$.

To prove condition (c), suppose that f is a band-limited image whose transform $F(u) \equiv 0$ for $|u| \geq b$, with $F(u) = 0$ in the interval $(-b, b)$ at $u = \pm z_1, \pm z_2, \ldots, \pm z_M$, where $0 < |z_1| < |z_2| < \ldots |z_M| = b$. If $a_{2,g} = a_{2,f}$ for another image g, the transform G has the same zeros as F. The proof of (b) showed that, for some constant c, $PhaG(u) = PhaF(u) + 2\pi cu$ for all u in the interval $(-z_1, z_1)$. We need to show that, for the same c,

$$PhaG(u) = PhaF(u) + 2\pi cu + 2\pi N \tag{A1}$$

for every u at which $F(u)$ is nonvanishing, with N an arbitrary integer. Suppose first that $M = 2$, i.e., that there is only one zero z_1 below the cutoff b. If $b \leq 2z_1$, then for all $u \in (z_1, b)$ we have $u/2 < z_1$. If we write $u = u/2 + u/2$, Eq. (9) implies that Eq. (A1) holds for $z_1 < u < 2z_1$. If $b > 2z_1$, then at $2z_1$ we write $2z_1 = 2z_1 - \epsilon + \epsilon$, and Eq. (9) shows that Eq. (8) holds at $u = 2z_1$. Then Eq. (A1) can be extended to $2z_1 < u < \max\{3z_1, b\}$ by writing $u = 2z_1 + pz_1, 0 < p < 1$. If we continue in this way, Eq. (A1) can eventually be shown to hold for all u in the interval (z_1, b), so $G(u) = \exp(-2\pi cu)F(u)$ for all u.

Now suppose that M is arbitrary. The argument above shows that Eq. (A1) holds in the interval (z_1, z_2). We assume that it holds for (z_{n-1}, z_n) and show that this implies that it holds for (z_n, z_{n+1}). Pick an integer N large enough that $z_1/N < z_n - z_{n-1}$, and write $u = z_n - z_1/N + pz_1$, with $1/N < p < 1$. Then $z_n < u < z_n + z_1(1 - 1/N)$, $pz_1 \in (0, z_1)$, and $z_n - z_1/N \in (z_{n-1}, z_n)$, so Eq. (8) implies Eq. (A1). As p ranges over $(1/N, 1)$, Eq. (A1) is extended in this way to cover all u in $[z_n, z_n + (1 - 1/N)z_1)]$. If this interval does not include the next zero z_{n+1}, then writing $z_n + (1 - 1/N)z_1$ in the form $u - \epsilon + \epsilon$ shows that Eq. (A1) holds at that u, and we can extend the relationship upward from that point by writing $u = z_n + (1 - 1/N)z_1 + pz_1$, with $0 < p < 1$ and applying Eq. (8) again. Repeating this process will eventually extend Eq. (A1) over the entire interval (z_n, z_{n+1}). Thus $PhaG(u) = PhaF(u) + 2\pi cu + 2\pi N$ for every $0 \leq u < b$ at which $F(u)$ is not zero, and the same is true for $-b < u \leq 0$ because $PhaG$ is odd. So $G(u) = \exp(i2\pi cu)F(u)$ for all u, and $g(x) = f(x + c)$.

Finally, for condition (d), suppose that F has infinite support but at most a finite number of zeros in any finite interval and that $a_{2,g} = a_{2,f}$ for some g. Then, for any u, either $F(u) = G(u) = 0$, so that $G(U) = \exp(i2\pi cu)F(u)$ trivially, or $F(u) \neq 0$; and for $0 \leq u' < u$, $F(u')$ either never vanishes or vanishes at a finite set of points $z_1 < z_2 < \ldots < z_n < u$. The first case is equivalent to condition (b), and in the second case the induction argument used to prove condition (c) can be used to show that $PhaG(u) = PhaF(u) + 2\pi cu + 2\pi N$, where the constant c is the same for all u at which F is nonvanishing. So again $G(u) = \exp(i2\pi cu)F(u)$ for all u, and $g(x) = f(x + c)$.

ACKNOWLEDGMENTS

We thank A. Ahumada, A. W. Lohmann, M. D'Zmura, R. Kakarala, S. Klein, and C. Tyler for information and advice.

REFERENCES AND NOTES

1. C. L. Nikias and J. M. Mendl, eds., *Workshop on Higher-Order Spectral Analysis,* IEEE Catalog No. 89THO267-5 (Institute of Electrical and Electronics Engineers, New York, 1989).
2. D. R. Brillinger, "Some history of higher-order moments and spectra," in *Workshop on Higher-Order Spectral Analysis* C. L. Nikias, and J. M. Mendl, eds., IEEE Catalog No. 89THO267-5 (Institute of Electrical and Electronics Engineers, New York, 1989).
3. A. W. Lohmann and B. Wirnitzer, "Triple correlation," Proc. IEEE **72,** 889–901 (1984).
4. S. A. Klein and C. W. Tyler, "Phase discrimination of compound gratings: generalized autocorrelation analysis," J. Opt. Soc. Am. A **3,** 868–879 (1986).
5. J. Aczel and J. I. Yellott, Jr., "Problem P159," Aeq. Math. **14,** 228 (1976).
6. J. I. Yellott, Jr., "The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution," J. Math. Psychol. **15,** 109–144 (1977).
7. C. Rockwell and J. I. Yellott, Jr., "A note on equivalent Thurstone models," J. Math. Psychol. **19,** 65–71 (1978).
8. Z. Moszner, "Sur l'equation $f(x)f(y)f(-x-y) = g(x)g(y)g(-x-y)$," Aeq. Math. **21,** 20–32 (1980).
9. Interest in higher-order autocorrelations appears to have evolved more or less independently in time-series analysis (Ref. 2), optics (Ref. 3), vision research (Ref. 4), and probabilistic choice theory (our own starting point: Refs. 5–8). Much recent work is described in Ref. 1.
10. J. I. Yellott and G. Iverson, "Triple correlation and texture discrimination," Invest. Ophthalmol. Vis. Sci. **31,** 561 (1990).
11. H. Bartelt, A. W. Lohmann, and B. Wirnitzer, "Phase and amplitude recovery from bispectra," Appl. Opt. **23,** 3121–3129 (1984).
12. G. B. Giannakis, "Signal reconstruction from multiple correlations: frequency and time domain approaches," J. Opt. Soc. Am. A **6,** 682–697 (1989). R. Kakarala has pointed out to us that the continuous-time versions of the discrete infinite duration signals considered by Giannakis are covered by our Theorem 3 and thus are uniquely determined up to translation by their triple correlations.
13. D. R. Brillinger, "The identification of a particular nonlinear time series system," Biometrika **64,** 509–515 (1977).
14. K. S. Lii and N. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," Ann. Stat. **10,** 1195–1208 (1982).
15. T. Matsuoka and T. J. Ulrych, "Phase estimation using the bispectrum," Proc. IEEE **72,** 1403–1411 (1984).
16. J_1 here is the first-order Bessel function.
17. J. A. Shohat and J. D. Tamarkin, *The Problem of Moments* (American Mathematical Society, Providence, R. I., 1970).
18. N. I. Akhiezer, *The Classical Problem of Moments* (Hafner, New York, 1965).
19. This problem has subsequently been solved by R. Kakarala and G. J. Iverson ("Characterization of periodic functions by higher order spectra," submitted to J. Opt. Soc. Am. A). They show that, in both one and two dimensions, every periodic function that is square integrable over its period is uniquely determined up to translation by its set of ACF's of all orders.
20. This definition is meant to include generalized functions of the form $f(x) = \Sigma_n b_n \delta(x - x_n)$, where δ is the Dirac delta and the b_n are positive numbers. The theory thus encompasses images described by sequences of discrete values. When such a sequence is finite, one may have the option of treating it as a function with bounded support and defining its higher-order ACF's by Eq. (3) or as a infinitely extended periodic function (as envisaged by the discrete Fourier transform), in which case the higher-order ACF's are defined by Eq. (4). From a uniqueness standpoint, the bounded-support option seems to be generally preferable, since it guarantees that the image is uniquely determined by its triple correlation. The periodic option offers no such guarantee: it is easy to construct pairs of discrete periodic images that are not translations of each other but that have the same triple correlation (e.g., sampled versions of the periodic functions in Fig. 2).
21. W. Feller, *An Introduction To Probability Theory and Its Applications* (Wiley, New York, 1966), Vol. II.
22. J. Aczel, *Lectures on Functional Equations* (Academic, New York, 1966).
23. E. Lukacs, *Characteristic Functions* (Charles Griffen, London, 1960).
24. The definition is meant to include generalized functions of the form $f(x, y) = \Sigma_n \Sigma_m b_{n,m} \delta(x - x_n, y - y_m)$. (See Note 20 above.)
25. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, San Francisco, Calif., 1968).