Fourier-based analysis of phase shifts in the superposition of periodic layers and their moiré effects

Isaac Amidror and Roger D. Hersch

Laboratoire de Systèmes Périphériques, Ecole Polytechnique
Fédérale de Lausanne, 1015 Lausanne, Switzerland

A Fourier-based approach is presented for the investigation of multilayer superpositions of periodic structures and their moiré effects. This approach fully explains the properties of the superposition of periodic layers and of their moiré effects, both in the spectral domain and in the image domain. We concentrate on showing how this approach provides also a full explanation of the various phenomena that occur because of phase shifts in one or more of the superposed layers. We show how such phase shifts influence the superposition as a whole and, in particular, how they affect each moiré in the superposition individually: We show that each moiré in the superposition undergoes a different shift, in its own main direction, whose size depends both on the moiré parameters and on the shifts of the individual layers. However, phase shifts in the individual layers do not necessarily lead to a solid shift of the whole superposition, and they may rather cause modifications in its microstructure. We demonstrate our results by several illustrative figures.
  Key words: moiré effect, superposition of periodic structures, superposition phase, phase shifts, Fourier analysis. © 1996 Optical Society of America

1. INTRODUCTION

The superposition of periodic structures (such as line gratings or dot screens) offers a wide range of interesting properties for exploration, starting from the overall structure of the superposition and culminating in the interesting and sometimes even spectacular moiré effects that may appear in the superposition. The importance of the superposition moiré phenomenon is clearly demonstrated by its vast number of applications in many different fields.[1–5] It is therefore important to understand this phenomenon and its various properties fully, along with the global properties of the layer superposition as a whole.

Although classical geometric or algebraic approaches can be used to explain the superposition moiré phenomenon and its geometric properties,[6,7] it has been shown that the best approach for exploring phenomena related to the superposition of periodic structures is the spectral approach, which is based on the Fourier theory.[8,9] Unlike the classical geometric and algebraic methods, this approach enables us to analyze properties not only in the original images and in their superposition but also in their spectral representations, and thus it offers a more profound insight into the problem and provides indispensable tools for exploring it. Moreover, the additional dimension offered by the impulse amplitudes in the spectrum (in addition to their geometric locations) also enables a quantitative analysis of the moiré intensity levels,[10] in addition to the qualitative geometric analysis of the moiré already offered by the earlier approaches.

One of the questions that were not fully investigated until now concerns the phase of the superposed layers; in fact, this question has been studied, by means of the classical approaches, only for the superposition of $m = 2$ layers.[1] It is important, however, to understand what

happens in the general m-layer superposition (and, particularly, to each of its moiré effects) when one or more of the superposed layers is being shifted.

Our main aim in the present paper is, therefore, to show how our spectral approach is particularly well adapted for investigating the influence of phase shifts in the superposition of any m periodic layers. We start, in Section 2, with a short review of our Fourier-based approach and its significant advantages, and we show how this approach explains the properties of the superposition and of the moiré effects generated thereby. In Section 3 we introduce the terminology and the basic notions related to phase shifts of 1-fold and 2-fold periodic functions. In Section 4 we analyze the effects of phase shifts in the superposition of any number of periodic layers. We allow the superposed layers to be freely or shifted on top of each other, with any desired phase difference. We show that each moiré in the superposition undergoes a different shift, and we derive the formula that gives the resulting shift for each of the moirés. Finally, we show that shifts in the individual superposed layers do not necessarily lead to a solid shift of the whole superposition but may rather cause modifications in its microstructure, and we give the conditions under which each of these cases occurs.

2. SPECTRAL APPROACH

The spectral approach is based on the duality between two-dimensional (2-D) images in the (x, y) plane and their 2-D spectra in the (u, v) frequency plane through the 2-D Fourier transform. Let us first review here the basic properties of the image types with which we are concerned and the fundamental notions and notations on which our spectral approach is based.

A. Properties of Our Images and Their Spectra

First, we deal here only with monochromatic (black-and-white) images. In this case each image can be represented in the image domain by a reflectance function, which assigns to any point (x, y) of the image a value between 0 and 1 representing its light reflectance: 0 for black (i.e., no reflected light), 1 for white (i.e., full light reflectance), and intermediate values for in-between shades. The images are superposed, for example, by overprinting. In the case of transparencies the reflectance function is replaced by a transmittance function defined in a similar way. Since the superposition of black and any other shade always gives black, this suggests a multiplicative model for the superposition of monochromatic images. Thus, when m monochromatic images are superposed, the reflectance of the resulting image is given by the product of the reflectance functions of the individual images:

$$r(x, y) = r_1(x, y) r_2(x, y) \cdots r_m(x, y). \tag{1}$$

According to the convolution theorem (Ref 11, p. 244) the Fourier transform of the product function is the convolution of the Fourier transforms of the individual functions:

$$R(u, v) = R_1(u, v) ** R_2(u, v) ** \cdots R_m(u, v). \tag{2}$$

Second, we are basically interested in periodic images defined on the continuous (x, y) plane, such as line gratings and dot screens, and their superpositions. This implies that the spectrum of the image on the (u, v) plane is not a continuous one but rather consists of impulses, corresponding to the frequencies that appear in the Fourier series decomposition of the image (Ref. 11, p. 204). A strong impulse in the spectrum indicates a pronounced periodic component in the original image at the frequency and the direction of that impulse. In the case of a 1-fold periodic image, such as a line grating, the spectrum consists of a one-dimensional (1-D) comb of impulses through the origin; in the case of a 2-fold periodic image the spectrum is a 2-D nailbed of impulses through the origin.

Each impulse in the 2-D spectrum is characterized by three main properties: its label (which is its index in the Fourier series development), its geometric location, and its amplitude (see Fig. 1). In terms of the original image the geometric location of an impulse in the spectrum determines the frequency f and the direction θ of the corresponding periodic component in the image, and the amplitude of the impulse represents the intensity of that periodic component in the image. (The impulse amplitude may be complex if the original image is not symmetric about the origin.)

However, the question of whether or not an impulse in the spectrum represents a visible periodic component in the image strongly depends on properties of the human visual system. The fact that the eye cannot distinguish fine details above a certain frequency (i.e., below a certain period) suggests that the human visual system model includes a low-pass filtering stage. This is a bidimensional bell-shaped filter [whose form is anisotropic, since it appears that the eye is less sensitive to small details in diagonal directions such as $45°$ (Ref. 12)]. However, for the sake of simplicity this low-pass filter can be approximated by the visibility circle, a circular step func-

tion around the spectrum origin whose radius represents the cutoff frequency (i.e., the threshold frequency beyond which fine detail is no longer detected by the eye). Obviously, its radius depends on several factors such as the contrast of the observed details, the viewing distance, and light conditions. If the frequencies of the original image elements are beyond the border of the visibility circle in the spectrum, the eye can no longer see them; but if a strong enough impulse in the spectrum of the image superposition falls inside the visibility circle, then a moiré effect becomes visible in the superposed image. (In fact, the visibility circle has a hole in its center, since very low frequencies cannot be seen, either.)

B. Spectrum Convolution and the Superposition Moirés

According to the convolution theorem [Eqs. (1) and (2)], when m line gratings are superposed in the image domain, the resulting spectrum is the convolution of their individual spectra. This convolution of combs (or nailbeds) can be seen as an operation in which frequency vectors from the individual spectra are added vectorially, while the corresponding impulse amplitudes are multiplied. More precisely, each impulse in the spectrum convolution is generated during the convolution process by the contribution of one impulse from each individual spectrum: its location is given by the sum of their frequency vectors, and its amplitude is given by the product of their amplitudes. This permits us to introduce an indexing method for denoting each of the impulses of the spectrum convolution in a unique, unambiguous way. The general impulse in the spectrum convolution will be denoted the $(k_1, k_2, \ldots, k_m)$ impulse, where m is the number of superposed gratings and each integer k_i is the index (harmonic), within the comb (the Fourier series) of the ith spectrum, of the impulse that this ith spectrum contributed to the impulse in question in the convolution. Using this formal notation, we have therefore the following result:

The geometric location of the general $(k_1, k_2, \ldots, k_m)$ impulse in the spectrum convolution is given by the vectorial sum (or linear combination)

$$\mathbf{f}_{k_1, k_2 \ldots, k_m} = k_1 \mathbf{f}_1 + k_2 \mathbf{f}_2 + \cdots + k_m \mathbf{f}_m, \tag{3}$$

and its amplitude is given by

$$a_{k_1, k_2, \ldots, k_m} = a_{k_1}^{(1)} a_{k_2}^{(2)} \cdots a_{k_m}^{(m)}, \tag{4}$$
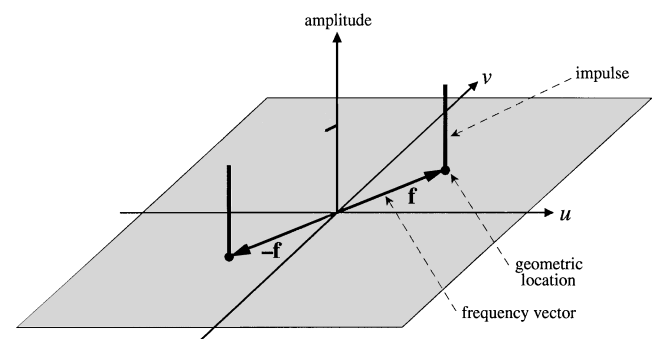

Fig. 1. Geometric location and amplitude of impulses in the 2-D spectrum. To each impulse is attached its frequency vector, which points to the geometric location of the impulse in the spectrum plane (u, v).

where $\mathbf{f}_i$ denotes the frequency vector of the fundamental impulse in the spectrum of the ith grating and $k_i\mathbf{f}_i$ and $a_{k_i}^{(i)}$ are, respectively, the frequency vector and the amplitude of the k_ith harmonic impulse in the spectrum of the ith grating.

The vectorial sum of Eq. (3) can also be written in terms of its Cartesian components. If f_i are the frequencies of the m original gratings and θ_i are the angles that they form with the positive horizontal axis, then the coordinates (f_u, f_v) of the $(k_1, k_2, \ldots, k_m)$ impulse in the spectrum convolution are given by

$$f_{u\,k_1,k_2,\ldots,k_m} = k_1f_1 \cos\theta_1 + k_2f_2 \cos\theta_2 + \cdots$$
$$+ k_mf_m \cos\theta_m,$$

$$f_{v\,k_1,k_2,\ldots,k_m} = k_1f_1 \sin\theta_1 + k_2f_2 \sin\theta_2 + \cdots$$
$$+ k_mf_m \sin\theta_m. \qquad (5)$$

Therefore the frequency, the period, and the angle of the considered impulse (and of the moiré that it represents) are given by the length and the direction of the vector $\mathbf{f}_{k_1,k_2,\ldots,k_m}$ as follows:

$$f = \sqrt{f_u{}^2 + f_v{}^2}, \qquad T_M = 1/f, \qquad \varphi_M = \arctan(f_v/f_u).$$
$$\qquad (6)$$

Note that in the special case of $m = 2$ gratings, when a moiré effect occurs because of the $(1, -1)$ impulse in the convolution, Eqs. (5) and (6) are reduced to the familiar

geometrically obtained formulas of the period and the angle of the moiré effect between two gratings[6]:

$$T_M = \frac{T_1T_2}{(T_1{}^2 + T_2{}^2 - 2T_1T_2 \cos\alpha)^{1/2}},$$

$$\sin\varphi_M = \frac{T_1 \sin\alpha}{(T_1{}^2 + T_2{}^2 - 2T_1T_2 \cos\alpha)^{1/2}} \qquad (7)$$

(where T_1 and T_2 are the periods of the two original images and α is the angle difference between them, $\theta_2 - \theta_1$). When $T_1 = T_2$, this is further simplified into the well-known formulas[13]

$$T_M = \frac{T}{2 \sin(\alpha/2)}, \qquad \varphi_M = 90° - \alpha/2. \qquad (8)$$

Let us now say a word about the notations used for the superposition moirés. As we have seen, a $(k_1, k_2, \ldots, k_m)$ impulse of the spectrum convolution that falls close to the spectrum origin, inside the visibility circle, represents a moiré effect in the superposed image (see Figs. 2 and 3). We call the m-grating moiré whose fundamental impulse is the $(k_1, k_2, \ldots, k_m)$ impulse in the spectrum convolution a $(k_1, k_2, \ldots, k_m)$ moiré; the highest absolute value in the index list is called the order of the moiré. Note that in the case of doubly periodic images, such as in dot screens, each superposed image contributes two frequency vectors to the spectrum, so that in Eqs. (3)–(5) m represents twice the number of superposed images.
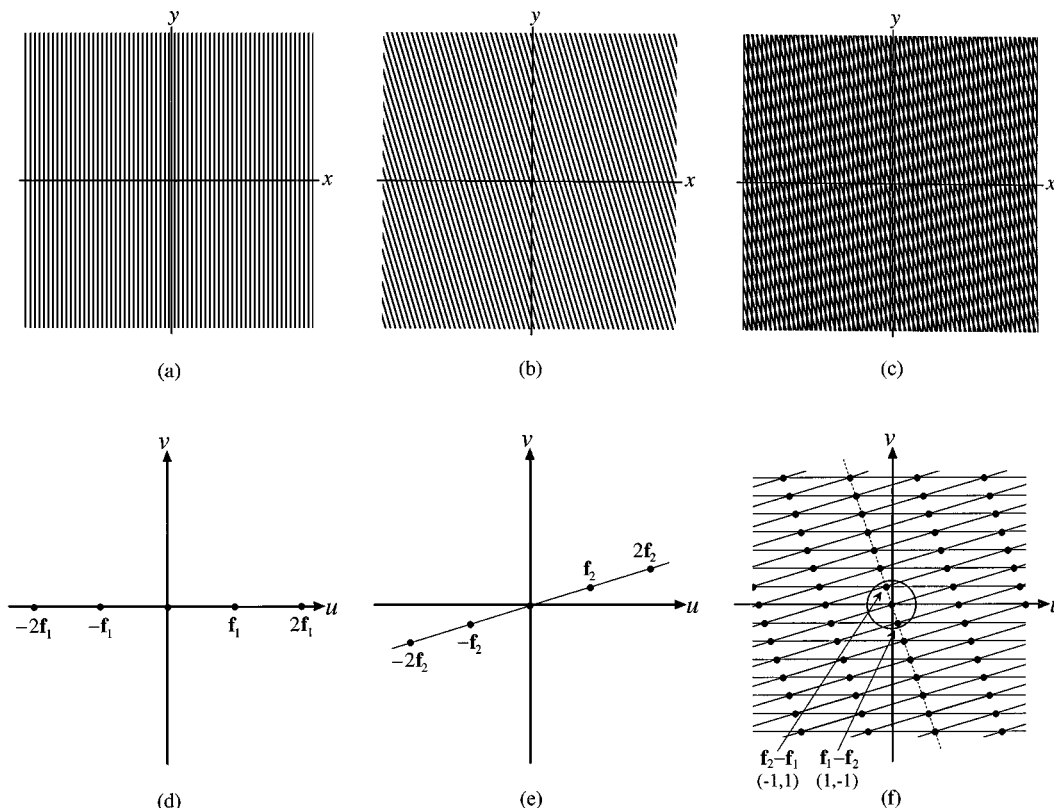

Fig. 2.   Binary gratings (a) and (b) and their superposition (c) in the image domain; their respective spectra are the infinite combs shown in (d) and (e) and their convolution (f).   Only impulse locations are shown in the spectra, and not their amplitudes.   The circle in the center of the spectrum (f) represents the visibility circle.   It contains the impulse pair whose frequency vectors are $\mathbf{f}_1 - \mathbf{f}_2$ and $\mathbf{f}_2 - \mathbf{f}_1$ and whose indices are $(1, -1)$ and $(-1, 1)$; this is the fundamental impulse pair of the $(1, -1)$ moiré seen in (c).   The dotted line in (f) shows the infinite comb of impulses that represents this moiré.
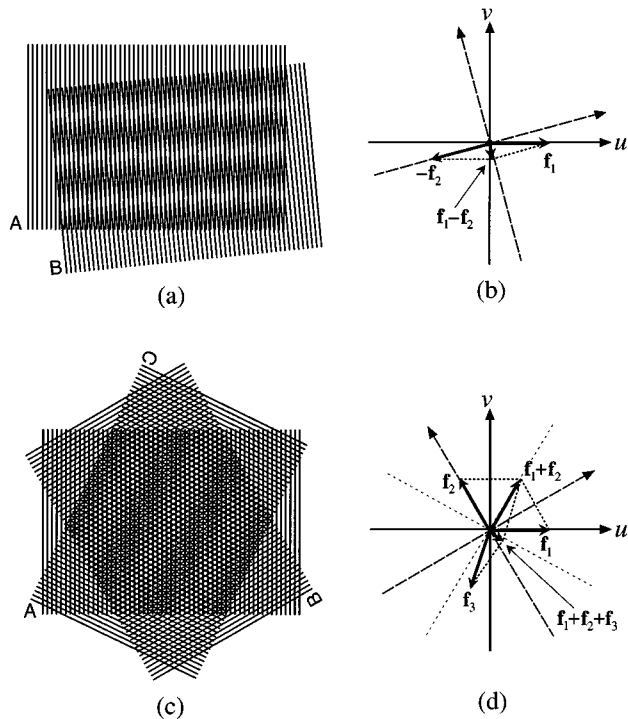
Fig. 3.   (a) The superposition of two identical gratings at a small angle difference gives rise to a $(1, -1)$ moiré. The spectral interpretation of (a) is shown in the vector diagram (b); compare with Fig. 2(f), which shows also impulses of higher harmonics. (c) The superposition of three identical gratings with angle differences slightly away from 120° gives a significant $(1, 1, 1)$ moiré; its vector diagram is shown in (d).

The spectrum of the superposition of periodic images, namely, the convolution of their original nailbed spectra, consists of a forest of impulses (with real or complex amplitudes, depending on the symmetry properties in the image domain). It has been shown[10] that the infinite impulse cluster in the spectrum convolution, which is centered on the spectrum origin and whose fundamental impulse is $(k_1, \ldots, k_m)$, represents in the spectrum the $(k_1, \ldots, k_m)$ moiré effect generated in the superposition. And indeed, by extracting from the spectrum convolution only this impulse cluster and taking its inverse Fourier transform, one obtains, back in the image domain, the isolated contribution of the moiré in question to the superposition, i.e., the moiré profile form (intensity levels).

C.   Advantages of the Spectral Approach

The spectral approach presented above proves fruitful in the investigation of the superposition of periodic layers and their resulting moiré effects. The most appealing features of the spectral approach include the following points:

1.   It provides a means for labeling and identifying each of the possible moiré effects in the m-layer superposition individually. The fact that every moiré is represented in the spectral domain by its fundamental impulse in the spectrum convolution (or, more precisely, by the impulse cluster that it spans there) means that each moiré in the superposition has its own identity or index notation: the $(k_1, \ldots, k_m)$ moiré.

2.   The spectrum convolution contains all the information about each moiré: the period and the angle of the moiré are given by the geometric location of its fundamental impulse, and its profile form is given by the impulse amplitudes of the fundamental impulse and its higher harmonics (the moiré cluster). This makes possible a full quantitative analysis of the intensity levels of each moiré, in addition to the qualitative geometric analysis of the moiré, which is already offered by the previous approaches.

3.   And most important, since the spectrum of the superposition contains *simultaneously* all the impulses that may represent moiré effects in the given superposition, it offers an overall, panoramic view of all the different moirés of various orders that are simultaneously present in the same layer superposition.[9]

4.   Moreover, this approach permits us to see how changes in the original superposed layers influence the spectrum. This enables us, in particular, to trace dynamically in the spectral domain the development of each of the moirés and to identify at any moment which of them are indeed visible.

5.   Our approach provides an easy explanation for all multiple-layer moirés, including the more complex cases in which the geometric analysis may become too complicated. In our approach all moirés of all orders are treated equally, and there is no longer any need to deal first with simple moirés, then with moirés of moirés, etc.

3.   PHASE TERMINOLOGY

The aim of the present paper is to analyze what happens in the superposition (and particularly to the moiré effects) when we change the relative phase between the superposed layers. In the following sections we will see how our spectral approach allows one, in a simple and elegant way, to cope with the question of the phases in the superposition of any m periodic layers. However, before proceeding to questions concerning the phases in the superposition and in its eventual moirés (Section 4), we will first define here the required terminology: first for the 1-D case and then for the 2-D case.

A.   Phase of a Periodic Function

Let $p(x, y)$ [or, in short, $p(\mathbf{x})$] be a 2-fold periodic (= skew-periodic) function with fundamental frequency vectors $\mathbf{f}_1$, $\mathbf{f}_2$ (see Fig. 11 in Appendix A). It is well known from the Fourier theory that $p(x, y)$ (if it satisfies some mathematical conditions, which are met by all our cases of interest) can be represented in the form of a 2-fold Fourier series[14]:

$$p(\mathbf{x}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \exp[i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{x}], \quad (9)$$

with the Fourier series coefficients

$$c_{m,n} = \frac{1}{A} \iint_A p(\mathbf{x})\exp[-i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{x}]d\mathbf{x}. \quad (10)$$

The spectrum of $p(\mathbf{x})$ is an oblique impulse nailbed, whose (m, n)th impulse has the frequency $m\mathbf{f}_1 + n\mathbf{f}_2$ and the complex amplitude $c_{m,n}$ [where $\mathbf{f}_1 = (u_1, v_1)$ and $\mathbf{f}_2 = (u_2, v_2)$ are the fundamental frequency vectors of the nailbed]:

$$P(\mathbf{u}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \delta[\mathbf{u} - (m\mathbf{f}_1 + n\mathbf{f}_2)]. \qquad (11)$$

Assume now that we shift $p(x, y)$ in the image domain by the vector $\mathbf{a} = (a, b)$, namely by a units in the x direction and b units in the y direction. We obtain for the shifted periodic function $p(x - a, y - b)$ the following expression:

$$p(\mathbf{x} - \mathbf{a}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \exp[i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2)$$
$$\cdot (\mathbf{x} - \mathbf{a})]$$
$$= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp[-i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{a}]$$
$$\times c_{m,n} \exp[i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{x}]. \qquad (12)$$

This is, in fact, a particular case of the general 2-D shift theorem (Ref. 11, p. 244), which is applied here to the case of 2-D periodic functions, i.e., to discrete (nailbed) spectra. The general 2-D shift theorem says that if the spectrum of a function $f(x, y)$ is $F(u, v)$, then the spectrum of the shifted function $f(x - a, y - b)$ is $\exp[-i2\pi(ua + vb)]F(u, v)$. In other words, a shift of $\mathbf{a} = (a, b)$ in the image domain multiplies the spectrum at each frequency $\mathbf{f} = (u, v)$ by the factor $\exp(-i2\pi\mathbf{f} \cdot \mathbf{a})$. And indeed, Eq. (12) shows that in our case, in which the spectrum is a nailbed, each (m, n) impulse in the spectrum, whose frequency is $\mathbf{f} = m\mathbf{f}_1 + n\mathbf{f}_2$, is multiplied by the factor $\exp[-i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{a}]$.

It follows from the shift theorem as a corollary that the phase increment generated in the spectrum as a result of a shift of $\mathbf{a}$ in the image domain is a linear function of the frequency, meaning that the increment in the phase spectrum that is due to the shift has the form of a continuous linear plane, whose slopes are determined by $\mathbf{a} = (a, b)$:

$$\phi(u, v) = \mathbf{f} \cdot \mathbf{a} = ua + vb. \qquad (13)$$

In our case, however, the spectrum of $p(\mathbf{x})$ is defined only in the points of the nailbed, i.e., at the frequencies of all the (m, n)th impulses: $(u, v) = m(u_1, v_1) + n(u_2, v_2)$, or, in vector notation, $\mathbf{f} = m\mathbf{f}_1 + n\mathbf{f}_2$. The phase increment generated at the (m, n)th impulse in the spectrum as a result of the shift of $\mathbf{a}$ in the image domain is

$$\phi_{m,n} = (m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{a}$$
$$= m\mathbf{f}_1 \cdot \mathbf{a} + n\mathbf{f}_2 \cdot \mathbf{a} \qquad (14)$$

(where $\mathbf{f}_1 \cdot \mathbf{a}$ and $\mathbf{f}_2 \cdot \mathbf{a}$ are constant numbers), which is simply the restriction of the linear plane of Eq. (13) to the points of our nailbed.

As we can see, a shift of the function $p(\mathbf{x})$ in the image domain influences only in the spectral domain the phases of the impulses; the impulse locations in the spectrum (as well as the absolute values of their amplitudes) are not influenced by the shift.

Expression (12) can be rewritten in yet another form as follows:

$$p(\mathbf{x} - \mathbf{a}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \exp[i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{x}$$
$$- i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{a}] \qquad (15)$$

or

$$p(\mathbf{x} - \mathbf{a}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \exp[i2\pi(m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{x}$$
$$- i2\pi\phi_{m,n}], \qquad (16)$$

where $\phi_{m,n} = (m\mathbf{f}_1 + n\mathbf{f}_2) \cdot \mathbf{a}$ is the phase component of the (m, n) impulse in the spectrum. We can see, therefore, that the phase of the function $p(x, y)$ may be expressed in the Fourier series representation of $p(x, y)$ either explicitly in the exponential part, as in Eq. (16), or implicitly, incorporated in the complex Fourier coefficients $c_{m,n}$ (the impulse amplitudes). The connection between these two representations of the phase can be clearly seen from Eq. (12), where the exponential constant representing the phase can be incorporated either in the main exponential part of the expression or in the coefficient $c_{m,n}$ itself.

If the function in the image domain is symmetric with respect to a certain point (x, y), it is natural to shift the function and to choose its initial phase (or its in-phase position) as the position in which its center of symmetry coincides with the origin. In this position the impulse amplitudes in the spectrum are purely real (or, in terms of modulus and phase, their phase components are constantly 0). But when the function is shifted from this position (by a nonintegral number of periods), its impulse amplitudes in the spectrum receive an imaginary component (or, in terms of modulus and phase, their phase component becomes a linear function of the frequency, according to the shift theorem).

However, in the case of a function having no point of symmetry, we will arbitrarily choose a certain shift position of the function as its initial phase, and we will fix this position by inserting its phase components inside the coefficients $c_{m,n}$. Once this initial phase and its corresponding coefficients $c_{m,n}$ have been fixed, they will be kept frozen, and from that moment on any phase shifts relative to this initial state will be represented only in terms of a phase component in the exponential, without modifying the fixed coefficients $c_{m,n}$.

It should be noted, however, that since we are dealing only with real images, their spectrum is always Hermitian (Ref. 11, p. 15), which means that the amplitudes of the impulse twins (which are symmetrically located on both sides of the origin, as in Fig. 1) are always complex conjugates.

B. Phase Terminology for Periodic Functions in the One-Dimensional Case

Let us introduce now some notations and terms in connection with the phase. We will start, for the sake of simplicity, with the 1-D case. Assume that $p(x)$ is a periodic function (symmetric or not) of period T. Clearly, since $p(x)$ is periodic, $p(x - nT)$ is identical to $p(x)$ for any integer n, so that a shift of $p(x)$ by an integer multiple of the period T is indistinguishable from the unshifted function. This fact suggests that any shift of $p(x)$ by a should be considered as being composed of an integer number of periods T plus a residue t:

$$a = nT + t \qquad 0 \le t < T,$$

where only the residue t affects the phase of $p(x - a)$, i.e.,

$$p(x - a) = p(x - nT - t) = p(x - t).$$

Consequently, we may introduce here the following notational conventions:

1. The total distance by which $p(x)$ has been displaced will be called the shift or the displacement of $p(x)$ and denoted by a [as in $p(x - a)$].

2. The residue $t = a \bmod T$, where $0 \le t < T$, will be called the effective shift (or effective displacement) of $p(x)$.

The shift and the effective shift of the periodic function $p(x)$ may also be expressed in terms of periods rather than in terms of distance units:

3. The shift of $p(x)$ expressed as a number (integer or not) of periods T will be called the period phase of $p(x)$ and denoted by $\phi = a/T = fa$. We have, therefore, $a = \phi T$.

4. The residue $r = t/T = \phi \bmod 1$, a number between 0 and 1 denoting the fraction of T by which $p(x)$ has been effectively shifted, will be called the effective period phase (or, in short, the effective phase) of $p(x)$.

As for the term phase itself, we will continue using it in the most general sense, including in the context of complex numbers (e.g., the modulus and phase representation of the spectrum, the phase component of an impulse).

As we can see, the influence of a shift of a on the phase of the periodic function $p(x)$ can be expressed in terms of the period phase ϕ as follows:

$$p(x - a) = \sum_{n=-\infty}^{\infty} c_n \exp[i2\pi nf(x - a)]$$
$$= \sum_{n=-\infty}^{\infty} c_n \exp(i2\pi nfx - i2\pi nfa)$$
$$= \sum_{n=-\infty}^{\infty} c_n \exp(i2\pi nfx - i2\pi n\phi). \qquad (17)$$

The phase component of the nth impulse in the spectrum of $p(x - a)$ is, therefore, $\phi_n = n\phi$.

C. Phase Terminology for 1-Fold Periodic Functions in the Two-Dimensional Case

We proceed now to the 2-D case and start with the simplest 2-D case—that of a 1-fold periodic function in an arbitrary direction in the (x, y) plane (such as a rotated grating; see Fig. 4). The 1-D terms that we defined above can be generalized into this 2-D case by considering them as vector quantities in the (x, y) plane.

Assume that $p(x, y)$ [or, in short, $p(\mathbf{x})$] is a 1-fold periodic function (symmetric or not) whose fundamental period vector is $\mathbf{T}$. Note that the fundamental period vector $\mathbf{T}$ of $p(\mathbf{x})$ is the shortest nonzero vector attached to the origin in the (x, y) plane that satisfies $p(\mathbf{x} + \mathbf{T}) = p(\mathbf{x})$. This shortest period vector is located on the x' axis, i.e., in the main direction of $p(\mathbf{x})$. Clearly, shifting $p(\mathbf{x})$ in the y' direction (i.e., perpendicularly to its main direction, x') has no effect, since $p(\mathbf{x})$ is constant in that direction. Therefore the effect of any arbitrary shift $\mathbf{a}'$ of $p(\mathbf{x})$ [that is, of moving the origin of $p(\mathbf{x})$ to the point $\mathbf{a}'$ in the (x, y) plane] is equivalent to a shift of $p(\mathbf{x})$ by $\mathbf{a} = \mathrm{proj}(\mathbf{a}')_{\mathbf{T}}$, the orthogonal projection of $\mathbf{a}'$ on the period vector $\mathbf{T}$ or on the x' axis [= the vector component of $\mathbf{a}'$ in the main direction of $p(\mathbf{x})$]. Therefore we will consider from now on only shifts of $p(\mathbf{x})$ along its main direction, and any other shift $\mathbf{a}'$ will be represented by its component $\mathbf{a}$ in that direction (see Fig. 4).

Since $p(\mathbf{x})$ is periodic with the fundamental period vector $\mathbf{T}$, $p(\mathbf{x} - n\mathbf{T})$ is identical to $p(\mathbf{x})$ for any integer n, so that a shift of $p(\mathbf{x})$ by an integer multiple of the vector period $\mathbf{T}$ is indistinguishable from the unshifted function. This fact suggests that any shift of $p(\mathbf{x})$ by $\mathbf{a}$ should be considered as being composed of an integer number of period vectors $\mathbf{T}$ plus a vectorial residue $\mathbf{t}$ (note that the vectors $\mathbf{T}$ and $\mathbf{a}$ are collinear, since both of them are located on the x' axis; therefore the vector $\mathbf{t}$ is also collinear with them):

$$\mathbf{a} = n\mathbf{T} + \mathbf{t}, \qquad \mathbf{t} = r\mathbf{T} \qquad 0 \le r < 1,$$

where only the residue $\mathbf{t}$ affects the phase of $p(\mathbf{x} - \mathbf{a})$, i.e.,

$$p(\mathbf{x} - \mathbf{a}) = p(\mathbf{x} - n\mathbf{T} - \mathbf{t}) = p(\mathbf{x} - \mathbf{t}).$$

Consequently, we may introduce here the following notational conventions as an extension of the 1-D case (Subsection 3.B):

1. The total distance by which $p(\mathbf{x})$ has been displaced will be called the shift or the displacement of $p(\mathbf{x})$ and denoted by $\mathbf{a}$ [as in $p(\mathbf{x} - \mathbf{a})$].

2. The residue $\mathbf{t} = \mathbf{a} \bmod \mathbf{T}$, where $\mathbf{0} \le \mathbf{t} < \mathbf{T}$ (meaning: $\mathbf{t} = r\mathbf{T}$, $0 \le r < 1$), will be called the effective shift (or effective displacement) of $p(\mathbf{x})$.

The shift and the effective shift of the periodic function $p(\mathbf{x})$ may also be expressed in terms of periods rather than in terms of distance units:

3. The shift of $p(\mathbf{x})$ expressed as a number (integer or not) of periods $\mathbf{T}$ will be called the period phase of $p(\mathbf{x})$ and denoted by $\phi = |\mathbf{a}|/|\mathbf{T}|$. We have, therefore, $\mathbf{a} = \phi\mathbf{T}$. Remembering that $\mathbf{T} \cdot \mathbf{T}^{-1} = 1$ [see Eq. (A5) in Appendix A], we multiply both sides (in the sense of scalar product) by $\mathbf{T}^{-1}$, and hence we obtain $\mathbf{a} \cdot \mathbf{T}^{-1} = \phi$, which can be expressed, as an extension of the 1-D case, as

$$\phi = \mathbf{f} \cdot \mathbf{a} \qquad (18)$$

[where $\mathbf{f} = \mathbf{T}^{-1}$ is the fundamental frequency vector of the 1-fold periodic function $p(\mathbf{x})$; see Eq. (A6) in Appendix A].

4. The residue $r = |\mathbf{t}|/|\mathbf{T}| = \phi \bmod 1$, a number between 0 and 1 denoting the fraction of $\mathbf{T}$ by which $p(\mathbf{x})$ has been effectively shifted, will be called the effective period phase (or, in short, the effective phase) of $p(\mathbf{x})$.
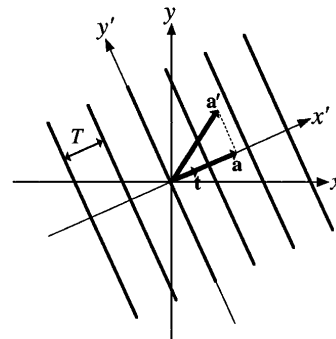

Fig. 4. Schematic plot of a 1-fold periodic function $p(\mathbf{x})$ in the image domain. A shift of $\mathbf{a}'$ from the origin is equivalent to a shift of $\mathbf{a}$ or to an effective shift of $\mathbf{t}$.

As in the 1-D case, we will continue to use the term phase itself in the general sense, whenever there is no need to specify which of the particular terms is meant.

As we can see, the influence of a shift of $\mathbf{a}$ on the phase of the 1-fold periodic function $p(\mathbf{x})$ can be expressed in terms of the period phase ϕ as follows:

$$
\begin{aligned}
p(\mathbf{x} - \mathbf{a}) &= \sum_{n=-\infty}^{\infty} c_n \exp[i2\pi n\mathbf{f} \cdot (\mathbf{x} - \mathbf{a})] \\
&= \sum_{n=-\infty}^{\infty} c_n \exp(i2\pi n\mathbf{f} \cdot \mathbf{x} - i2\pi n\mathbf{f} \cdot \mathbf{a}) \\
&= \sum_{n=-\infty}^{\infty} c_n \exp(i2\pi n\mathbf{f} \cdot \mathbf{x} - i2\pi n\phi). \quad (19)
\end{aligned}
$$

The phase component of the nth impulse in the spectrum of $p(\mathbf{x} - \mathbf{a})$ is, therefore, $\phi_n = n\phi$.

D. Phase Terminology for the General Two-Dimensional Case: 2-Fold Periodic Functions

We proceed now to the generalization of the above phase terminology to the general 2-D case. Assume that $p(\mathbf{x})$ is a 2-fold periodic function (symmetric or not) whose fundamental frequency vectors in the spectrum are $\mathbf{f}_1$ and $\mathbf{f}_2$. As shown in Appendix A, the periodic properties of $p(\mathbf{x})$ can be expressed in the image domain in two alternative ways: either in terms of the classical period vectors $\mathbf{P}_1$ and $\mathbf{P}_2$, which are defined by

$$
\mathbf{P}_i \cdot \mathbf{f}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},
$$

or in terms of step vectors $\mathbf{T}_1$ and $\mathbf{T}_2$, which we define as vectors collinear to $\mathbf{f}_1$ and $\mathbf{f}_2$ with reciprocal lengths:

$$
\mathbf{T}_i = \frac{1}{|\mathbf{f}_i|} \frac{\mathbf{f}_i}{|\mathbf{f}_i|} = \frac{1}{|\mathbf{f}_i|^2} \mathbf{f}_i .
$$

Accordingly, the phase in the 2-D case can also be presented in two ways (see Fig. 5). However, because of the clear advantages of the step-vector notation, as shown in Appendix A, we give here the phase terminology in terms of step vectors only.

Let $p(\mathbf{x})$ be a 2-fold periodic function as above (see Fig. 5). Consider the fundamental period parallelogram of $p(\mathbf{x})$ defined by the vectors $\mathbf{P}_1$, $\mathbf{P}_2$, and let $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ be two 1-fold periodic functions whose periods are defined by the borders of this parallelogram (see the solid and dotted gratings in Fig. 5). If the given 2-fold periodic function $p(\mathbf{x})$ is separable, i.e., if it can be presented as a product of two 1-D functions, then $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ are simply the two 1-D periodic components of $p(\mathbf{x})$: $p(\mathbf{x}) = p_1(\mathbf{x})p_2(\mathbf{x})$. If, however, $p(\mathbf{x})$ is inseparable (like a screen with circular dots), we choose $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ to be the two gratings defined by the parallelogram borders (as in Fig. 5), and instead of the original function $p(\mathbf{x})$ we consider the separable superposition $p'(\mathbf{x}) = p_1(\mathbf{x})p_2(\mathbf{x})$; we will call $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ the virtual gratings of $p(\mathbf{x})$. Note that the period vectors $\mathbf{T}_1$ and $\mathbf{T}_2$ of the 1-fold periodic functions $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ are the step vectors of the 2-fold periodic function $p(\mathbf{x})$.

Clearly, $p'(\mathbf{x})$ and $p(\mathbf{x})$ have the same periods and the same phases, and, in fact, they differ in their Fourier series representations by only their coefficients $c_{m,n}$ [see

Eq. (9)]. Therefore, as far as periodicity and phases are concerned, $p(\mathbf{x})$ and $p'(\mathbf{x})$ are equivalent, and we may represent the inseparable function $p(\mathbf{x})$ by the separable function $p'(\mathbf{x})$. For the sake of simplicity we will proceed with the discussion below using $p'(\mathbf{x})$; in the case of a separable function $p(\mathbf{x})$ we may simply replace $p'(\mathbf{x})$ by $p(\mathbf{x})$ itself.

Now the idea is to look at each of the two superposed 1-fold periodic virtual gratings separately: Obviously, any shift of $p'(\mathbf{x})$ by $\mathbf{a}$ can be considered as a simultaneous shift of both $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ by $\mathbf{a}$. However, as we have seen in Subsection 3.C, a shift of $p_1(\mathbf{x})$ by $\mathbf{a}$ is equivalent to a shift of $p_1(\mathbf{x})$ by $\mathbf{a}_1 = \mathrm{proj}(\mathbf{a})_{\mathbf{T}_1}$; and a shift of $p_2(\mathbf{x})$ by $\mathbf{a}$ is equivalent to a shift of $p_2(\mathbf{x})$ by $\mathbf{a}_2 = \mathrm{proj}(\mathbf{a})_{\mathbf{T}_2}$. For example, in the case shown in Fig. 5 we have $\mathbf{a}_1 = 2.5\mathbf{T}_1$ and $\mathbf{a}_2 = 3.5\mathbf{T}_2$.

We see that the shift of $p'(\mathbf{x})$ by $\mathbf{a}$ is identical to a shift of the grating $p_1(\mathbf{x})$ by $\mathbf{a}_1 = \phi_1\mathbf{T}_1$ and a shift of the grating $p_2(\mathbf{x})$ by $\mathbf{a}_2 = \phi_2\mathbf{T}_2$. This is equivalent, in turn, to an effective shift of the grating $p_1(\mathbf{x})$ by $\mathbf{t}_1 = r_1\mathbf{T}_1$ and an effective shift of the grating $p_2(\mathbf{x})$ by $\mathbf{t}_2 = r_2\mathbf{T}_2$ (where r_1 and r_2 represent the fractional parts of ϕ_1 and ϕ_2).

$$\mathbf{a} = 2.5\mathbf{P}_1 + 3.5\mathbf{P}_2$$

(a)

$$\mathbf{a}_1 = \mathrm{proj}(\mathbf{a})_{\mathbf{T}_1} = 2.5\mathbf{T}_1$$
$$\mathbf{a}_2 = \mathrm{proj}(\mathbf{a})_{\mathbf{T}_2} = 3.5\mathbf{T}_2$$
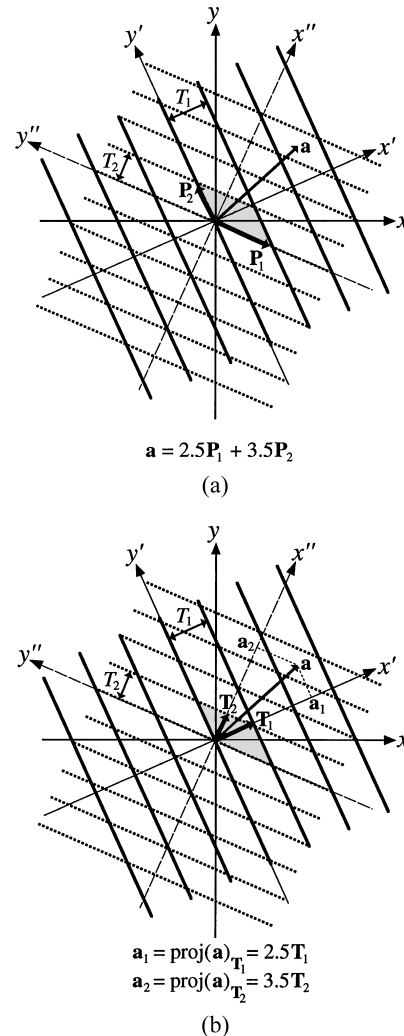
(b)

Fig. 5. Schematic plot of the unshifted 2-fold periodic function $p(\mathbf{x})$ in the image domain: (a) a vector $\mathbf{a}$, expressed in terms of the period vectors $\mathbf{P}_1$, $\mathbf{P}_2$; (b) the same vector $\mathbf{a}$, expressed in terms of the step vectors $\mathbf{T}_1$, $\mathbf{T}_2$.

As we can see, we reduced here the phase of the 2-fold periodic function into the phases of its two 1-fold periodic virtual gratings, and hence we can simply use here the phase terminology of Subsection 3.C, applied to each of the virtual gratings separately. Consequently, we may introduce here the following notational conventions as an extension of the 1-fold periodic case:

1. The total distance by which $p(\mathbf{x})$ has been displaced is called the shift or the displacement of $p(\mathbf{x})$ and denoted by $\mathbf{a}$ [as in $p(\mathbf{x} - \mathbf{a})$]. As we have seen, this is equivalent to shifts of $\mathbf{a}_1$ and $\mathbf{a}_2$ in the virtual gratings $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$, where $\mathbf{a}_1$ and $\mathbf{a}_2$ are the projections of the vector $\mathbf{a}$ in the directions of $\mathbf{T}_1$ and $\mathbf{T}_2$.

2. The residue $\mathbf{t}$, called the effective shift (or effective displacement) of $p(\mathbf{x})$, can be expressed in a similar way by $\mathbf{t}_1$ and $\mathbf{t}_2$, where $\mathbf{t}_1$ and $\mathbf{t}_2$ are the effective shifts of the virtual gratings $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ (see convention 2 of Subsection 3.C): $\mathbf{t}_1 = \mathbf{a}_1 \bmod \mathbf{T}_1$, $\mathbf{t}_2 = \mathbf{a}_2 \bmod \mathbf{T}_2$. In other words, $\mathbf{t}_1$ and $\mathbf{t}_2$ are the projections of the vector $\mathbf{t}$ in the directions of $\mathbf{T}_1$ and $\mathbf{T}_2$.

The shift and the effective shift of $p(\mathbf{x})$ may also be expressed in terms of periods rather than in terms of distance units:

3. The shift of $p(\mathbf{x})$ expressed as a pair of numbers (integer or not) representing the number of periods $\mathbf{T}_1$, $\mathbf{T}_2$ of $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ is called the period phase of $p(\mathbf{x})$ and denoted by $\boldsymbol{\phi} = (\phi_1, \phi_2)$, where ϕ_1 and ϕ_2 are the period phases of each of the two virtual gratings $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ (see convention 3 of Subsection 3.C).

4. The residue $\mathbf{r} = (r_1, r_2)$, a pair of numbers between 0 and 1 denoting the fractions of $\mathbf{T}_1$, $\mathbf{T}_2$ by which $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ have been effectively shifted, is called the effective period phase (or, in short, the effective phase) of $p(\mathbf{x})$.

Now the influence of a shift $\mathbf{a}$ on the phase of the 2-fold periodic function $p'(\mathbf{x})$ can be expressed in terms of the period phases ϕ_1 and ϕ_2 of $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ as follows:

$$p'(\mathbf{x} - \mathbf{a}) = p_1(\mathbf{x} - \mathbf{a}_1) p_2(\mathbf{x} - \mathbf{a}_2)$$

$$= \left\{ \sum_{m=-\infty}^{\infty} c_m^{(1)} \exp[i2\pi m \mathbf{f}_1 \cdot (\mathbf{x} - \mathbf{a}_1)] \right\}$$

$$\times \left\{ \sum_{n=-\infty}^{\infty} c_n^{(2)} \exp[i2\pi n \mathbf{f}_2 \cdot (\mathbf{x} - \mathbf{a}_2)] \right\}$$

$$= \left[\sum_{m=-\infty}^{\infty} c_m^{(1)} \exp(i2\pi m \mathbf{f}_1 \cdot \mathbf{x} - i2\pi m \mathbf{f}_1 \cdot \mathbf{a}_1) \right]$$

$$\times \left[\sum_{n=-\infty}^{\infty} c_n^{(2)} \exp(i2\pi n \mathbf{f}_2 \cdot \mathbf{x} - i2\pi n \mathbf{f}_2 \cdot \mathbf{a}_2) \right]$$

$$= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_m^{(1)} c_n^{(2)} \exp[i2\pi(m \mathbf{f}_1 + n \mathbf{f}_2) \cdot \mathbf{x}$$
$$- i2\pi(m \mathbf{f}_1 \cdot \mathbf{a}_1 + n \mathbf{f}_2 \cdot \mathbf{a}_2)]$$

$$= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_m^{(1)} c_n^{(2)} \exp[i2\pi(m \mathbf{f}_1 + n \mathbf{f}_2) \cdot \mathbf{x}$$
$$- i2\pi(m\phi_1 + n\phi_2)], \tag{20}$$

where $\phi_1 = \mathbf{f}_1 \cdot \mathbf{a}_1 = |\mathbf{a}_1|/|\mathbf{T}_1|$ and $\phi_2 = \mathbf{f}_2 \cdot \mathbf{a}_2 = |\mathbf{a}_2|/|\mathbf{T}_2|$

are the period phases of the two shifted 1-fold periodic functions $p_1(\mathbf{x} - \mathbf{a}_1)$ and $p_2(\mathbf{x} - \mathbf{a}_2)$ [see Eq. (18)].

Once we have obtained the phase of the separable 2-fold periodic function $p'(\mathbf{x}) = p_1(\mathbf{x}) p_2(\mathbf{x})$ after the shift, all that we need to do in order to obtain the equivalent expression for the original, inseparable function $p(\mathbf{x})$ itself is to replace in Eq. (20) the Fourier coefficients $c_m^{(1)} c_n^{(2)}$ that belong to $p'(\mathbf{x})$ by the inseparable Fourier coefficients $c_{m,n}$ of the original function $p(\mathbf{x})$. We obtain

$$p(\mathbf{x} - \mathbf{a}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \exp[i2\pi(m \mathbf{f}_1 + n \mathbf{f}_2) \cdot \mathbf{x}$$
$$- i2\pi(m\phi_1 + n\phi_2)]. \tag{21}$$

Comparing this expression with Eq. (16), which was directly obtained by the shift theorem, we see that the phase component of the (m, n)th impulse in the spectrum of $p(\mathbf{x} - \mathbf{a})$ is $\phi_{m,n} = m\phi_1 + n\phi_2$, where ϕ_1, ϕ_2 are the period phases of the two shifted virtual gratings $p_1(\mathbf{x} - \mathbf{a}_1)$ and $p_2(\mathbf{x} - \mathbf{a}_2)$. This is the 2-D equivalent of $\phi_n = n\phi$ in the 1-D case and in the 1-fold periodic case (see at the end of Subsections 3.B and 3.C).

Note that the above generalization of the phase concepts from scalar quantities in the 1-D case (Subsection 3.B) to the equivalent vectorial quantities in the 2-D case, with scalar product replacing the simple number multiplication (e.g., $\phi = \mathbf{f} \cdot \mathbf{a}$ replacing $\phi = fa$), was made possible thanks to the step-vector approach (see Appendix A).

4. PHASES IN THE SUPERPOSITION

When we superpose two or more periodic functions in their initial phase, we obtain the initial phase of the superposition (and of its moirés). Note that the initial phase is simply distinguished by the absence of a phase component in the exponent of the Fourier series (the initial phase being incorporated in the coefficients $c_{m,n}$). What happens now if we shift one or more of the superposed layers, i.e., if we modify their phase with respect to their initial phase? It should be noted, first, that since phase shifts in any individual layer do not influence the impulse locations in its own spectrum, they do not influence the impulse locations in the spectrum convolution, either. Therefore we obtain the following result:

Result 1: Phase shifts in any of the superposed layers have no influence on the support of the spectrum of the superposition and on its algebraic structure. In particular, such phase shifts have no influence on the spectrum support of any moiré generated in the superposition, meaning that the period and angle of the moiré remain unchanged. ∎

Let us now see in detail the influence of shifts in the individual superposed layers on their superposition.

Let $p_1(\mathbf{x}), \ldots, p_m(\mathbf{x})$ be m 1-fold periodic functions given in their initial phases, and let $p(\mathbf{x}) = p_1(\mathbf{x}) \cdots p_m(\mathbf{x})$ be their superposition (in its initial phase). Note that each pair of 1-fold periodic functions may represent in the superposition one 2-fold periodic function, such as a dot screen. We have, therefore,

$$p(\mathbf{x}) = p_1(\mathbf{x}) \cdots p_m(\mathbf{x})$$

$$= \left[\sum_{k_1=-\infty}^{\infty} c_{k_1}^{(1)} \exp(i2\pi k_1 \mathbf{f}_1 \cdot \mathbf{x}) \right] \times \cdots$$

$$\times \left[\sum_{k_m=-\infty}^{\infty} c_{k_m}^{(m)} \exp(i2\pi k_m \mathbf{f}_m \cdot \mathbf{x}) \right]$$

$$= \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_m=-\infty}^{\infty} c_{k_1}^{(1)} \cdots c_{k_m}^{(m)} \exp[i2\pi(k_1\mathbf{f}_1 + \cdots$$

$$+ k_m\mathbf{f}_m) \cdot \mathbf{x}]. \tag{22}$$

The $(k_1, \ldots, k_m)$ moiré extracted from this superposition (see the end of Subsection 2.B) contains only the impulses of the $(k_1, \ldots, k_m)$ comb, i.e., the $(nk_1, \ldots, nk_m)$ impulses for any integer n and is therefore given by

$$m_{k_1,\ldots,k_m}(\mathbf{x}) = \sum_{n=-\infty}^{\infty} c_{nk_1}^{(1)} \cdots c_{nk_m}^{(m)} \exp[i2\pi(k_1\mathbf{f}_1 + \cdots$$

$$+ k_m\mathbf{f}_m) \cdot \mathbf{x}]. \tag{23}$$

Assume now that each of the layers $p_i(\mathbf{x})$ is shifted from its initial position by $\mathbf{a}_i$ (for any layer i that remains unshifted we simply take $\mathbf{a}_i = \mathbf{0}$). We have, therefore,

$$p_1(\mathbf{x} - \mathbf{a}_1) \cdots p_m(\mathbf{x} - \mathbf{a}_m)$$

$$= \left\{ \sum_{k_1=-\infty}^{\infty} c_{k_1}^{(1)} \exp[i2\pi k_1 \mathbf{f}_1 \cdot (\mathbf{x} - \mathbf{a}_1)] \right\} \times \cdots$$

$$\times \left\{ \sum_{k_m=-\infty}^{\infty} c_{k_m}^{(m)} \exp[i2\pi k_m \mathbf{f}_m \cdot (\mathbf{x} - \mathbf{a}_m)] \right\}$$

$$= \left\{ \sum_{k_1=-\infty}^{\infty} c_{k_1}^{(1)} \exp[i2\pi k_1 \mathbf{f}_1 \cdot \mathbf{x} - i2\pi k_1 \mathbf{f}_1 \cdot \mathbf{a}_1] \right\} \times \cdots$$

$$\times \left[\sum_{k_m=-\infty}^{\infty} c_{k_m}^{(m)} \exp(i2\pi k_m \mathbf{f}_m \cdot \mathbf{x} - i2\pi k_m \mathbf{f}_m \cdot \mathbf{a}_m) \right]$$

$$= \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_m=-\infty}^{\infty} c_{k_1}^{(1)} \cdots c_{k_m}^{(m)} \exp[i2\pi(k_1\mathbf{f}_1 + \cdots$$

$$+ k_m\mathbf{f}_m) \cdot \mathbf{x} - i2\pi(k_1\mathbf{f}_1 \cdot \mathbf{a}_1 + \cdots + k_m\mathbf{f}_m \cdot \mathbf{a}_m)]$$

$$= \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_m=-\infty}^{\infty} c_{k_1}^{(1)} \cdots c_{k_m}^{(m)} \exp[i2\pi(k_1\mathbf{f}_1 + \cdots$$

$$+ k_m\mathbf{f}_m) \cdot \mathbf{x} - i2\pi(k_1\phi_1 + \cdots + k_m\phi_m)]. \tag{24}$$
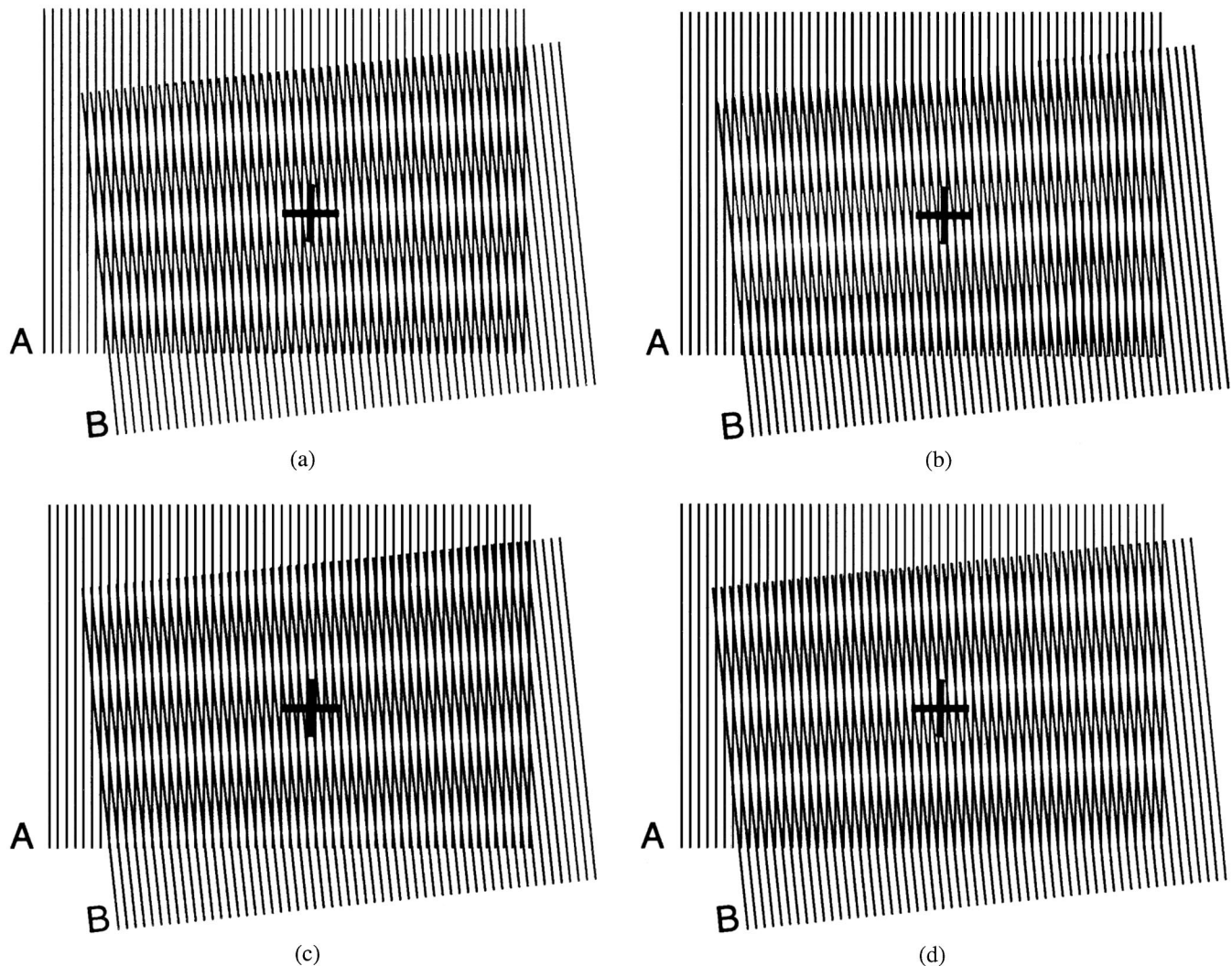

Fig. 6. Two-grating $(1, -1)$ moiré of Figs. 3(a) and 3(b) and its phase shifts that are due to a shift in grating A. The origin of each image is indicated by a cross. (a) Both gratings and the moiré are in their initial phase. (b) Grating A is shifted by 1/4 period to the right; the moiré is consequently shifted 1/4 moiré period downward. (c) Grating A is shifted by 1/2 period to the right; the moiré is consequently shifted 1/2 moiré period downward. (d) Grating A is shifted by 3/4 period to the right; the moiré is consequently shifted 3/4 moiré period downward.
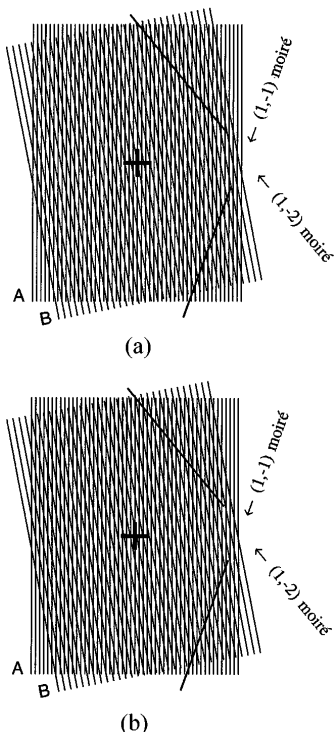
(a)

(b)

Fig. 7.　$(1, -1)$ and $(1, -2)$ moirés between two gratings and their phase shifts that are due to a shift in grating B.　(a) Both gratings and the moirés are in their initial phase.　(b) Grating B is shifted by $1/4$ period to the right; consequently, the $(1, -1)$ moiré is shifted by $1/4$ of its period, while the $(1, -2)$ moiré is shifted by $1/2$ of its period (note that the cross and the lines indicating the moiré directions have not been moved).

It should be noted that if an inseparable 2-fold periodic function (such as a screen with circular dots) takes part in the superposition, instead of two of the 1-fold periodic functions, then the only difference in Eq. (24) is that two of the coefficients, $c_{k_i}^{(i)}$ and $c_{k_j}^{(j)}$, will be replaced by a single inseparable coefficient, $c_{k_i,k_j}^{(i,j)}$.

The $(k_1, \ldots, k_m)$ moiré extracted from this superposition contains only the impulses of the $(k_1, \ldots, k_m)$ comb and is therefore given by

$$\sum_{n=-\infty}^{\infty} c_{nk_1}^{(1)} \cdots c_{nk_m}^{(m)} \exp[i2\pi n(k_1\mathbf{f}_1 + \cdots$$
$$+ k_m\mathbf{f}_m) \cdot \mathbf{x} - i2\pi n(k_1\phi_1 + \cdots + k_m\phi_m)]. \quad (25)$$

Comparing this expression with Eq. (23), we see that the exponent here is simply shifted by $i2\pi n(k_1\phi_1 + \cdots + k_m\phi_m)$, where the expression in parentheses is a constant number.　This suggests, by the shift theorem, that expression (25) simply represents a shifted version of Eq. (23); therefore, if we denote by $\mathbf{b}_{k_1,\ldots,k_m}$ the unknown shift that the $(k_1, \ldots, k_m)$ moiré $m_{k_1,\ldots,k_m}(\mathbf{x})$ itself has undergone, we can rewrite expression (25) as

$$m_{k_1,\ldots,k_m}(\mathbf{x} - \mathbf{b}_{k_1,\ldots,k_m})$$
$$= \sum_{n=-\infty}^{\infty} c_{nk_1}^{(1)} \cdots c_{nk_m}^{(m)} \exp[i2\pi n(k_1\mathbf{f}_1 + \cdots$$
$$+ k_m\mathbf{f}_m) \cdot \mathbf{x} - i2\pi n(k_1\phi_1 + \cdots + k_m\phi_m)]. \quad (26)$$

On the other hand, according to Eq. (12) the effect of shifting $m_{k_1,\ldots,k_m}(\mathbf{x})$ [Eq. (23)] by $\mathbf{b}_{k_1,\ldots,k_m}$ is

$$m_{k_1,\ldots,k_m}(\mathbf{x} - \mathbf{b}_{k_1,\ldots,k_m})$$
$$= \sum_{n=-\infty}^{\infty} c_{nk_1}^{(1)} \cdots c_{nk_m}^{(m)} \exp[i2\pi n(k_1\mathbf{f}_1 + \cdots$$
$$+ k_m\mathbf{f}_m) \cdot (\mathbf{x} - \mathbf{b}_{k_1,\ldots,k_m})]$$
$$= \sum_{n=-\infty}^{\infty} c_{nk_1}^{(1)} \cdots c_{nk_m}^{(m)} \exp[i2\pi n(k_1\mathbf{f}_1 + \cdots$$
$$+ k_m\mathbf{f}_m) \cdot \mathbf{x}$$
$$- i2\pi n(k_1\mathbf{f}_1 + \cdots + k_m\mathbf{f}_m) \cdot \mathbf{b}_{k_1,\ldots,k_m}]. \quad (27)$$

By comparing the exponents in Eqs. (26) and (27), we find that for every n, i.e., for every term in the summation, we have

$$(k_1\mathbf{f}_1 + \cdots + k_m\mathbf{f}_m) \cdot \mathbf{b}_{k_1,\ldots,k_m} = k_1\phi_1 + \cdots + k_m\phi_m.$$

Since $k_1\mathbf{f}_1 + \cdots + k_m\mathbf{f}_m$ is the frequency vector $\mathbf{f}_{k_1,\ldots,k_m}$ of the $(k_1, \ldots, k_m)$ impulse [= the fundamental impulse of the $(k_1, \ldots, k_m)$ moiré], we obtain

$$\mathbf{f}_{k_1,\ldots,k_m} \cdot \mathbf{b}_{k_1,\ldots,k_m} = k_1\phi_1 + \cdots + k_m\phi_m. \quad (28)$$

Remembering now that $\mathbf{f} \cdot \mathbf{f}^{-1} = 1$ (see Appendix A), we multiply both sides (in the sense of scalar product) by $(\mathbf{f}_{k_1,\ldots,k_m})^{-1}$:

$$\mathbf{b}_{k_1,\ldots,k_m} = (\mathbf{f}_{k_1,\ldots,k_m})^{-1}(k_1\phi_1 + \cdots + k_m\phi_m),$$
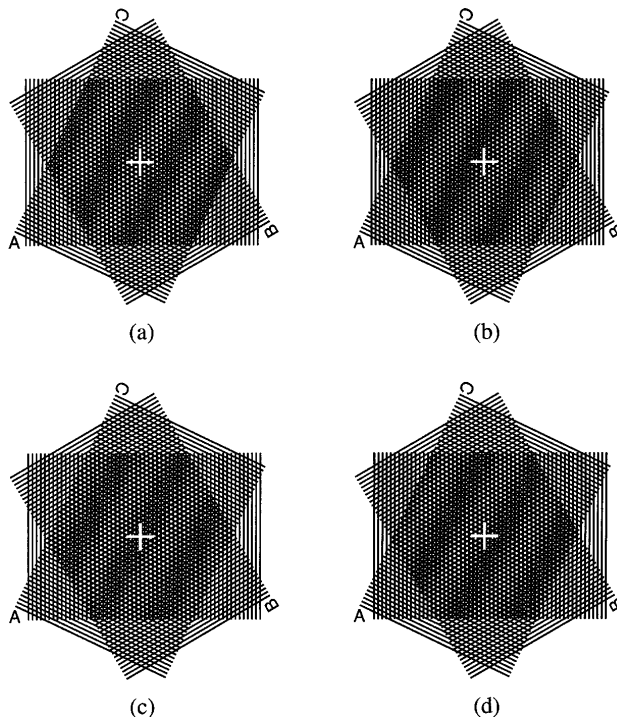

(a)

(b)

(c)

(d)

Fig. 8.　Three-grating $(1, 1, 1)$ moiré of Figs. 3(c) and 3(d) and its phase shifts that are due to a shift in grating A.　(a) The three gratings and the moiré are in their initial phase.　(b) Grating A is shifted by $1/4$ period to the right; consequently, the moiré is shifted $1/4$ moiré period in the main direction of the moiré. (c) Grating A is shifted by $1/2$ period to the right; consequently, the moiré is shifted $1/2$ moiré period in the main direction of the moiré.　(d) Grating A is shifted by $3/4$ period to the right; consequently, the moiré is shifted $3/4$ moiré period in the main direction of the moiré.

which by Eq. (A6) becomes

$$\mathbf{b}_{k_1,\ldots,k_m} = \mathbf{T}_{k_1,\ldots,k_m}(k_1\phi_1 + \cdots + k_m\phi_m), \qquad (29)$$

where $\mathbf{T}_{k_1,\ldots,k_m}$ is the step vector of the $(k_1, \ldots, k_m)$ moiré in the image domain and $\phi_i = |\mathbf{a}_i|/|\mathbf{T}_i|$. The shift that the $(k_1, \ldots, k_m)$ moiré undergoes is given, therefore, by the period of the moiré multiplied by the scalar coefficient $k_1\phi_1 + \cdots + k_m\phi_m$, which is a linear combination of the period phases $\phi_1, \ldots, \phi_m$ of the individual superposed gratings. Note the similarity with $\mathbf{a} = \phi\mathbf{T}$ in the case of a single grating (convention 3 of Subsection 3.C).

Furthermore, denoting the period phase of the $(k_1, \ldots, k_m)$ impulse [the period phase of the $(k_1, \ldots, k_m)$ moiré] by $\phi_{k_1,\ldots,k_m}$, we obtain from Eqs. (18) and (28) the expression

$$\phi_{k_1,\ldots,k_m} = k_1\phi_1 + \cdots + k_m\phi_m. \qquad (30)$$

We obtain, therefore, the following important result:

Result 2: When the original m superposed layers $p_1(\mathbf{x}), \ldots, p_m(\mathbf{x})$ are shifted by $\mathbf{a}_1, \ldots, \mathbf{a}_m$, respectively, each of the $(k_1, \ldots, k_m)$ moirés generated in the superposition merely undergoes a shift of $\mathbf{b}_{k_1,\ldots,k_m}$, which is a multiple of its step vector by a scalar, as given by Eq. (29). The size and the direction of the shift are therefore different for each of the $(k_1, \ldots, k_m)$ moirés in the superposition. ∎

Example 1: In the particular case of the $(1, -1)$ moiré between $m = 2$ gratings we have

$$\mathbf{b}_{1,-1} = \mathbf{T}_{1,-1}(\phi_1 - \phi_2), \qquad (31)$$

$$\phi_{1,-1} = \phi_1 - \phi_2. \qquad (32)$$

Figure 6 illustrates what happens to the phase of the $(1, -1)$ moiré between two superposed gratings when we shift the first grating (A) along its main direction (= horizontally) while keeping the second grating (B) at its initial phase (so that $\mathbf{a}_2 = \mathbf{0}$): When $\mathbf{a}_1 = \mathbf{0}$, we obtain $\mathbf{b}_{1,-1} = \mathbf{0}$ and the moiré is in phase; and as we gradually increase the shift $\mathbf{a}_1$ from $\mathbf{0}$ to $\mathbf{T}_1 = \mathbf{f}_1^{-1}$ (one full period of the first grating), the moiré is gradually shifted along its main direction from $\mathbf{b}_{1,-1} = \mathbf{0}$ to $\mathbf{b}_{1,-1} = \mathbf{T}_{1,-1}$ (namely, one full period of the moiré). Note that the shift of the moiré is larger than the shift of the grating (in the same proportion as their periods) and is in a different direction (the direction of the moiré period).

Example 2: Consider the $(1, -1)$ and $(1, -2)$ moirés of Fig. 7(a), between $m = 2$ gratings. If we shift grating B along its main direction by $1/4$ period ($\mathbf{a}_2 = \mathbf{T}_2/4$) while keeping grating A at its initial phase ($\mathbf{a}_1 = \mathbf{0}$), we find from Eq. (29) that the $(1, -1)$ moiré is shifted by $1/4$ moiré period ($\mathbf{b}_{1,-1} = -\mathbf{T}_{1,-1}/4$), while the $(1, -2)$ moiré is shifted by $1/2$ moiré period ($\mathbf{b}_{1,-2} = -\mathbf{T}_{1,-2}/2$, since $k_2 = 2$). This is indeed confirmed in Fig. 7(b).

An example with the three-grating $(1, 1, 1)$ moiré is shown in Fig. 8, and an example with the $(1, 0, -1, 0)$ and $(0, 1, 0, -1)$ moirés between $m = 4$ gratings is given in Fig. 9.

It is important to note, however, that the overall superposition in the image domain after the original layers have been shifted by $\mathbf{a}_1, \ldots, \mathbf{a}_m$ is not, in general, a shifted version (a rigid motion) of the original superposition, and the microstructure of the superposition may be modified (see, for example, Fig. 10). In fact, we can say the following:

It is clear that every shift $\mathbf{a}$ of the superposition $p_1(\mathbf{x}) \cdots p_m(\mathbf{x})$ can be seen as a combination of individual shifts $\mathbf{a}_1, \ldots, \mathbf{a}_m$ of the gratings $p_1(\mathbf{x}), \ldots, p_m(\mathbf{x})$ along their main directions. However, the converse is not necessarily true: Let, for example, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ be the individual shifts of the gratings $p_1(\mathbf{x}), p_2(\mathbf{x}), p_3(\mathbf{x})$ that correspond to the shift $\mathbf{a}$ of the whole superposition $p_1(\mathbf{x})p_2(\mathbf{x})p_3(\mathbf{x})$. Obviously, $\mathbf{a}_1, \mathbf{a}_2,$ and $\mathbf{a}_3$ are linearly dependent, since once $\mathbf{a}_1$ and $\mathbf{a}_2$ [the components of $\mathbf{a}$ in the main directions of $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$] have been fixed, $\mathbf{a}_3$ must join them in order that the superposition be shifted as a whole. Therefore $\mathbf{a}_3$ is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$: $\mathbf{a}_3 = s_1\mathbf{a}_1 + s_2\mathbf{a}_2$ (with some fixed constants $s_1, s_2 \in \mathbb{R}$). However, if we choose $\mathbf{a}_3$ differently, the superposition $p_1(\mathbf{x})p_2(\mathbf{x})p_3(\mathbf{x})$ will not be shifted as a whole, but rather it will be modified. We have, therefore, the following result:

Result 3: Nontrivial shifts of $\mathbf{a}_1, \ldots, \mathbf{a}_m$ in the gratings $p_1(\mathbf{x}), \ldots, p_m(\mathbf{x})$ in their respective main directions (i.e., nonidentical shifts $\mathbf{a}_1, \ldots, \mathbf{a}_m$ that are not inte-
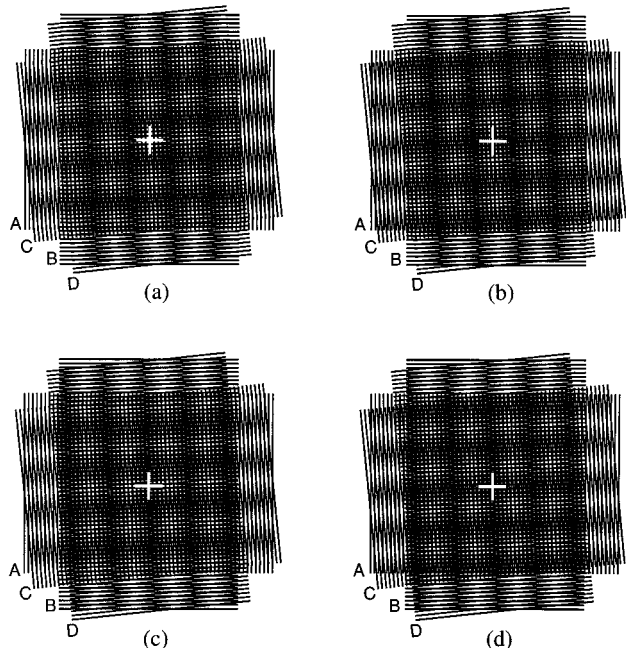

Fig. 9. Perpendicular $(1, 0, -1, 0)$ and $(0, 1, 0, -1)$ moirés between two grids (four gratings) and their phase shifts that are due to a shift in the first grid (gratings A and B). (a) Both grids and the two perpendicular moirés are in their initial phase. (b) The first grid (gratings A and B) is shifted by $1/2$ period to the right; consequently, the horizontal moiré is shifted $1/2$ moiré period downward. (c) The first grid (gratings A and B) is shifted by $1/2$ period upward; consequently, the vertical moiré is shifted $1/2$ moiré period to the right. (d) The first grid (gratings A and B) is shifted by $1/2$ period to the right and $1/2$ period upward; consequently, the horizontal moiré is shifted $1/2$ moiré period downward, and the vertical moiré is shifted $1/2$ period to the right.
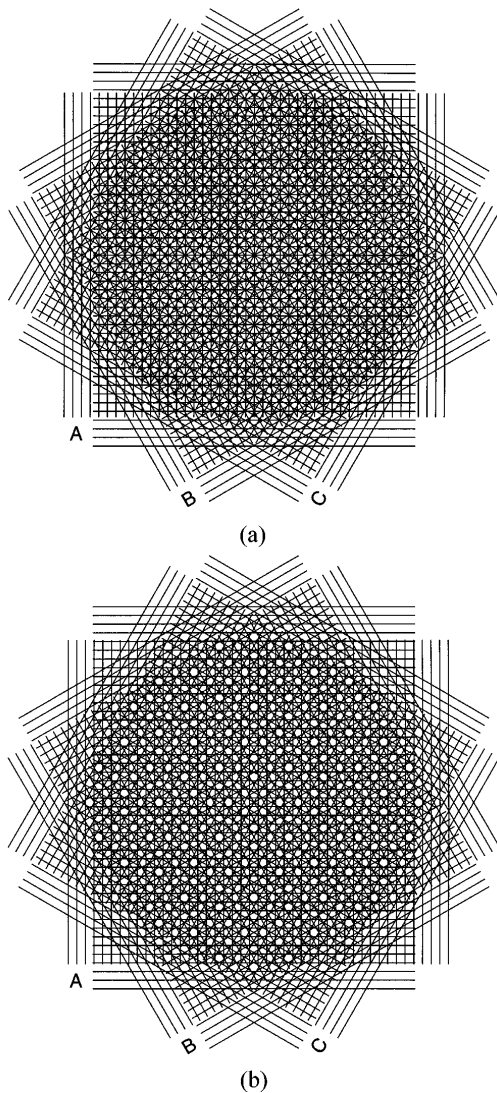
(a)

(b)

Fig. 10.   Magnification of the superposition of three grids (six gratings) with identical periods and equal angle differences of 30°. This is an example of an almost-periodic superposition. In (a) all the grids are superposed in their initial phase, while in (b) the grid A has been shifted by 1/2 period in its two primary directions; note the substantial change in the form of the microstructure that is due to this phase shift.

gral linear combinations of $\mathbf{T}_1, \ldots, \mathbf{T}_m$) do not, in general, correspond to a shift $\mathbf{a}$ of the whole superposition $p_1(\mathbf{x}) \cdots p_m(\mathbf{x})$. They do, however, if and only if (iff) the shifts $\mathbf{a}_i$ are projections of the same vector $\mathbf{a}$ on the main directions of $p_i(\mathbf{x})$ or, in other words, iff the normal lines defined by each of the shifts $\mathbf{a}_i$ perpendicularly to the main axis of $p_i(\mathbf{x})$ [Fig. 5(b)] meet in a single point $\mathbf{a}$. ∎

For example, in the case of $m > 2$ gratings, shifting just one layer (by a nontrivial shift) does not correspond to a shift of the whole superposition, and it rather modifies the superposition microstructure. Note that in the case of $m = 2$ gratings the conditions of Result 3 can be simplified as follows: Nontrivial shifts of $\mathbf{a}_1$ in $p_1(\mathbf{x})$ and of $\mathbf{a}_2$ in $p_2(\mathbf{x})$ in their respective main directions correspond to a shift $\mathbf{a}$ in the whole superposition $p_1(\mathbf{x})p_2(\mathbf{x})$ iff $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ (i.e., $\mathbf{T}_1$ and $\mathbf{T}_2$) are not collinear.

Note that Results 2 and 3 may seem at first sight con-

tradictory: Suppose that $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ are two superposed gratings (with noncollinear $\mathbf{T}_1, \mathbf{T}_2$). How can it be that shifts of $\mathbf{a}_1$ and $\mathbf{a}_2$ in $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ cause a shift of the whole superposition by $\mathbf{a}$, while, on the other hand, each of the different (k_1, k_2) moirés that are present in that superposition undergoes a different shift and in a different direction? In fact, as we can also see in Figs. 6 and 7, there exists no contradiction here: The whole superposition is indeed shifted by the same vector $\mathbf{a}$, but as we have seen in Subsection 3.C (see Fig. 4), this common shift is then translated into the proper language of each individual moiré by considering its projection on the moiré's own main direction.

The strength of Result 2 is that it remains valid in all cases: each of the $(k_1, \ldots, k_m)$ moirés in the superposition is individually shifted (without changing its period, angle, or profile form), even when the shifts of the individual gratings do not correspond to a rigid motion of the whole superposition but rather modify its microstructure.

5.   SUMMARY

The superposition of any number m of periodic layers (such as line gratings, dot screens, etc.) and the phenomena related thereto, such as the superposition moiré effects, can be fully explained by means of the spectral approach, which is based on the Fourier theory. After a short presentation of this approach we proceed, in the main part of the paper, to show how this approach allows one, in a simple and elegant way, to cope also with the various questions related to the phases between several superposed layers. We show how phase shifts in the superposed layers influence the superposition as a whole and how it affects each moiré in the superposition individually. The key point is that a shift of any periodic function in the image domain influences in the spectral domain only the phases of the impulses and not their geometric locations or the absolute value of their amplitudes. We show that nontrivial shifts of each of the individual superposed layers may result either in a global shift (rigid motion) of the whole superposition or in a modification of its microstructure. Regarding the moiré effects in the superposition, we show that shifts in one or more of the individual superposed layers do not influence the periods, the angles, and the profile shapes of any of the moirés. The only effect of such shifts on the moirés resides in a shift of each of the $(k_1, \ldots, k_m)$ moirés in the superposition. The size and the direction of the shift are different for each of the moirés in the superposition, and they are given by Eq. (29). This result holds in all cases, even when the individual layer shifts do not cause a global shift of the whole superposition but rather modify its microstructure.

APPENDIX A: PERIOD VECTORS P_i VERSUS STEP VECTORS T_i

In this appendix we introduce an alternative formalism for describing the relationship between periodicities in the image domain and the frequencies in the spectral domain, which is based on the new notion of step vectors rather than on the classical period vectors. The distinct advantages of this formalism in the case of periodic layer superpositions are discussed and illustrated.
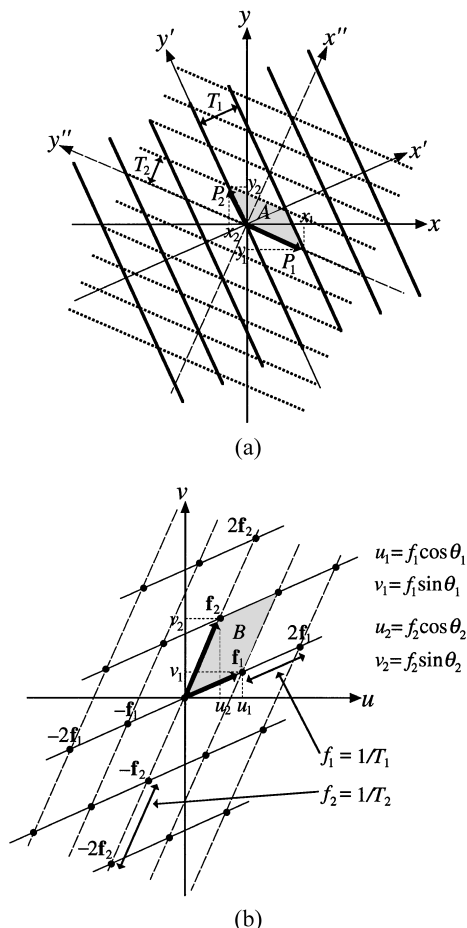
(a)

(b)

Fig. 11.   Schematic plot of (a) the 2-fold periodic (skew-periodic) function $p(x, y)$ in the image domain and (b) its skewed impulse nailed in the spectral domain.   The gray parallelogram A in the image domain represents a one-period element (tile) of $p(x, y)$. P_1 and P_2 are segments of this parallelogram that coincide with the period vectors $\mathbf{P}_1$ and $\mathbf{P}_2$.   It can be shown that the areas of parallelograms A and B are reciprocal: $\overline{B} = 1/\overline{A}$.

A function $p(x, y)$ is called 2-fold periodic if there exist two nonzero and noncollinear vectors $\mathbf{P}_1 = (x_1, y_1)$ and $\mathbf{P}_2 = (x_2, y_2)$ so that, for all $(x, y) \in \mathbb{R}^2$,

$$p(x + x_1, y + y_1) = p(x, y) \qquad p(x + x_2, y + y_2) = p(x, y).$$

The vectors $\mathbf{P}_1$ and $\mathbf{P}_2$ are called periods or period vectors of $p(x, y)$.   (As in the case of frequency vectors, we always consider period vectors as radius vectors emanating from the origin, and hence the period parallelograms that they define are attached to the origin.)   Note that $p(x, y)$ is completely determined from its values in the period parallelogram A defined by $\mathbf{P}_1$ and $\mathbf{P}_2$, which repeats itself through the (x, y) plane [see Fig. 11(a)].

It is therefore customary to describe the periodicity of a 2-fold periodic function $p(x, y)$ in $\mathbb{R}^2$ by means of two period vectors $\mathbf{P}_1$ and $\mathbf{P}_2$.   As is well known,[15] the period vectors $\mathbf{P}_1$, $\mathbf{P}_2$ of the function $p(x, y)$ are obtained from the fundamental frequency vectors $\mathbf{f}_1$, $\mathbf{f}_2$ in the spectrum of $p(x, y)$ [see Eq. (11)] by

$$\mathbf{P}_i \cdot \mathbf{f}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \tag{A1}$$

This can also be expressed in matrix notation, as follows (for n dimensions):

$$\begin{pmatrix} \mathbf{P}_1 \\ \vdots \\ \mathbf{P}_n \end{pmatrix}(\mathbf{f}_1, \ldots, \mathbf{f}_n) = \begin{bmatrix} \mathbf{P}_1 \cdot \mathbf{f}_1 & \cdots & \mathbf{P}_1 \cdot \mathbf{f}_n \\ \cdots & \cdots & \cdots \\ \mathbf{P}_n \cdot \mathbf{f}_1 & \cdots & \mathbf{P}_n \cdot \mathbf{f}_n \end{bmatrix}$$

$$= \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

(note that each vector $\mathbf{P}_i$ or $\mathbf{f}_i$ represents here an n-tuple of coordinates, so that we are dealing with $n \times n$ matrices).   It follows that if the matrix $(\mathbf{f}_1, \ldots, \mathbf{f}_n)$ is invertible, i.e., nonsingular (which is true iff the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are linearly independent over $\mathbb{R}$),[16] then

$$\begin{pmatrix} \mathbf{P}_1 \\ \vdots \\ \mathbf{P}_n \end{pmatrix} = (\mathbf{f}_1, \ldots, \mathbf{f}_n)^{-1}. \tag{A2}$$

However, the period-vector notation is valid only when the frequency vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are linearly independent. When the frequency vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are linearly dependent, the matrix on the right-hand side of Eq. (A2) is singular, i.e., noninvertible, and no corresponding $\mathbf{P}_i$ vectors exist in the image domain.   This means that the period vectors $\mathbf{P}_i$ are inappropriate for our study of the layer superpositions, because layer superpositions may well occur even with $\mathbf{f}_1, \ldots, \mathbf{f}_n$ linearly dependent.   (And, in fact, in the case of our 2-D spectrum this necessarily occurs whenever the number of superposed layers is larger than 2.)   Therefore, instead of the vectors $\mathbf{P}_i$, in the image domain we will consider only the periods T_i of the n individual gratings (see Fig. 11).   Note that a 2-fold periodic layer $p(x, y)$ in the superposition (say, a dot screen) will contribute two values T_i, which are the periods of the two virtual gratings defined by the borders of the fundamental period parallelogram of that layer [see the solid and dotted gratings in Fig. 11(a)].[17]   And indeed, even the classical moiré formulas, such as Eqs. (7) and (8), are based on only the grating periods T_i.   If we consider these grating periods as vectors $\mathbf{T}_i$ emanating from the origin of the image domain at the grating directions θ_i (see Fig. 12) and call them step vectors, we find that the relationship between each step vector $\mathbf{T}_i$ and its counterpart $\mathbf{f}_i$ in the spectral domain is straightforward:

1.   The vectors $\mathbf{T}_i$ and $\mathbf{f}_i$ (for every i) are collinear, i.e., they have the same angle θ_i.
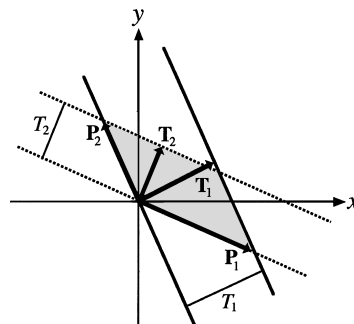2.   Their lengths are reciprocal, namely, for every i, $|\mathbf{T}_i| = 1/|\mathbf{f}_i|$.

Fig. 12.   Magnified view of the main period parallelogram of Fig. 11(a), showing the period vectors $\mathbf{P}_i$ and the step vectors $\mathbf{T}_i$.

This means that for every i we have $\mathbf{T}_i = (|\mathbf{T}_i|/|\mathbf{f}_i|)\mathbf{f}_i = (1/|\mathbf{f}_i|^2)\mathbf{f}_i$.   In other words, we obtain

$$
\begin{aligned}
\mathbf{T}_1 &= \frac{1}{|\mathbf{f}_1|^2}\,\mathbf{f}_1 \\
&\vdots \\
\mathbf{T}_n &= \frac{1}{|\mathbf{f}_n|^2}\,\mathbf{f}_n\,.
\end{aligned}
\tag{A3}
$$

As we can see, this relationship holds between any pair $\mathbf{T}_i$, $\mathbf{f}_i$ individually; and moreover, it exists in all cases and for any number n of superposed layers—even if $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are linearly dependent.

To formulate better the relationship between the vector pair $\mathbf{T}_i$, $\mathbf{f}_i$, we introduce the following definition:

Definition: For any vector $\mathbf{v} \neq 0$ the reciprocal vector of $\mathbf{v}$ (with respect to scalar product) is defined as

$$
\mathbf{v}^{-1} = \frac{1}{|\mathbf{v}|^2}\,\mathbf{v}\,.
\tag{A4}
$$

This definition requires a short explanation.   Although it is clear that the vector $\mathbf{v}^{-1}$ is reciprocal to $\mathbf{v}$ with respect to scalar product,

$$
\mathbf{v}^{-1}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{v}^{-1} = \frac{1}{|\mathbf{v}|^2}\,\mathbf{v}\cdot\mathbf{v} = 1\,,
\tag{A5}
$$

$\mathbf{v}^{-1}$ is not the unique vector with this property.   In fact, the locus of all the vectors $\mathbf{x}$ that satisfy $\mathbf{x}\cdot\mathbf{v} = 1$ consists (in the case of $\mathbb{R}^3$) of the whole plane perpendicular to the vector $\mathbf{v}$, whose distance from the origin (along the line spanned by $\mathbf{v}$) is $1/|\mathbf{v}|$.   Therefore the uniqueness of $\mathbf{v}^{-1}$ in this definition is obtained only through the requirement, implicit in Eq. (A4), that the reciprocal vector $\mathbf{v}^{-1}$ be collinear with the vector $\mathbf{v}$.

Using this definition and in view of relationships 1 and 2 above and Eq. (A3), we can now reformulate the relationship between the vector pair $\mathbf{T}_i$ and $\mathbf{f}_i$ as follows:

3.   For every i the vector $\mathbf{T}_i$ in the image domain is the reciprocal vector (with respect to scalar product) of the frequency vector $\mathbf{f}_i$ in the spectral domain:

$$
\mathbf{T}_i = \frac{1}{|\mathbf{f}_i|^2}\,\mathbf{f}_i = \mathbf{f}_i{}^{-1}\,.
\tag{A6}
$$

This is, in fact, the 2-D vectorial generalization of the relation $T = 1/f$ between period and frequency in the 1-D case.

Returning now to our vector comparison, the main difference between the step vectors $\mathbf{T}_i$ and the period vectors $\mathbf{P}_i$ (all of which subsist in the image domain) is that each vector $\mathbf{T}_i$ depends only on a single frequency vector $\mathbf{f}_i$, while, as we have seen in Eqs. (A1) and (A2), each of the vectors $\mathbf{P}_i$ depends on all of the n frequency vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$.   In the first case we speak about reciprocity (with respect to scalar product) between the individual vectors $\mathbf{T}_i$ and $\mathbf{f}_i$, but in the second case we speak about reciprocity between vector n-tuples, $\mathbf{P}_1, \ldots, \mathbf{P}_n$ and $\mathbf{f}_1, \ldots, \mathbf{f}_n$, or between the lattices $L_\mathbf{P}$ and $L_\mathbf{f}$ spanned by them.   And while the vectors $\mathbf{P}_i$ exist only when the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are linearly independent, the vectors $\mathbf{T}_i$, to the contrary, exist for any vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ with no restrictions, since every vector $\mathbf{T}_i$ is dependent only on its own counterpart $\mathbf{f}_i$, and no matrix inversion is involved.   Note that when the vector frequencies $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are all orthogonal to each other (in the n-dimensional spectrum), the vectors $\mathbf{T}_i$ and $\mathbf{P}_i$ ($i = 1, \ldots, n$) coincide in the image domain.   In particular, in any 1-fold periodic function the vectors $\mathbf{T}$ and $\mathbf{P}$ coincide.

Finally, it should be emphasized that in spite of the apparent symmetry between the frequency lattice $L_\mathbf{f}$ and the period lattice $L_\mathbf{P}$ (or between the frequency vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ and the period vectors $\mathbf{P}_1, \ldots, \mathbf{P}_n$) that is due to Eqs. (A1) and (A2), there exists a substantial difference between them: While the spectrum support $L_\mathbf{f}$ is a fundamental property of any superposition of n periodic functions, the period lattice $L_\mathbf{P}$ (and the period vectors $\mathbf{P}_1, \ldots, \mathbf{P}_n$) are only derived properties, and they exist only conditionally: iff the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are linearly independent (i.e., iff the superposition in the image domain is periodic in n dimensions).

REFERENCES AND NOTES

1. K. Patorski, *Handbook of the Moiré Fringe Technique* (Elsevier, Amsterdam, 1993).
2. O. Kafri and I. Glatt, *The Physics of Moiré Metrology* (Wiley, New York, 1989).
3. A. T. Shepherd, "25 years of moiré fringe measurement," Precis. Eng. **1,** 61–69 (1979).
4. H. Takasaki, "Moiré topography," Appl. Opt. **9,** 1467–1472 (1970).
5. P. S. Theocaris, *Moiré Fringes in Strain Analysis* (Pergamon, Oxford, 1969).
6. Y. Nishijima and G. Oster, "Moiré patterns: their application to refractive index and refractive index gradient measurements," J. Opt. Soc. Am. **54,** 1–5 (1964).
7. G. Oster, M. Wasserman, and C. Zwerling, "Theoretical interpretation of moiré patterns," J. Opt. Soc. Am. **54,** 169–175 (1964).
8. O. Bryngdahl, "Moiré: formation and interpretation," J. Opt. Soc. Am. **64,** 1287–1294 (1974).
9. I. Amidror, R. D. Hersch, and V. Ostromoukhov, "Spectral analysis and minimization of moiré patterns in colour separation," J. Electron. Imaging **3,** 295–317 (1994).
10. I. Amidror, "A generalized Fourier-based method for the analysis of 2D moiré envelope forms in screen superpositions," J. Mod. Opt. **41,** 1837–1862 (1994).
11. R. N. Bracewell, *The Fourier Transform and Its Applications*, 2nd ed. (McGraw-Hill, New York, 1986).
12. R. Ulichney, *Digital Halftoning* (MIT Press, Cambridge, Mass., 1988), pp. 79–84.
13. G. Oster, *The Science of Moiré Patterns*, 2nd ed. (Edmund Scientific, Barrington, N.J., 1969).
14. See, for example, A. Papoulis, *Systems and Transforms with Applications in Optics* (McGraw-Hill, New York, 1968), p. 117.
15. See, for example, A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, 2nd ed. (Academic, Orlando, Fla., 1982), Vol. 1, p. 75.
16. G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, 4th ed. (Macmillan, New York, 1977), pp. 237–238.
17. Although there exist different possible choices of fundamental frequency vectors $\mathbf{f}_1$, $\mathbf{f}_2$ for $p(x, y)$, each of these choices automatically determines a corresponding pair of fundamental period vectors $\mathbf{P}_1$, $\mathbf{P}_2$ [by Eq. (A1)], and hence it determines also the fundamental period parallelogram that they define and the corresponding virtual-grating periods $\mathbf{T}_1$, $\mathbf{T}_2$.   Note that all the different possible choices represent the same 2-D lattices $L_\mathbf{f}$ and $L_\mathbf{P}$ of $p(x, y)$.