

Journal of Electronic Imaging, 3(3), pages 295-317, 1994.

Spectral Analysis and Minimization of Moiré Patterns in Colour Separation

**Isaac Amidror,
Roger D. Hersch,
Victor Ostromoukhov**

Swiss Federal Institute of Technology (EPFL),
Peripheral Systems Laboratory (LSP),
1015 Lausanne, SWITZERLAND.

Abstract

Undesired moiré patterns may appear in colour printing for various reasons. One of the most important reasons is interference between the superposed halftone screens of the different primary colours, due to an improper alignment of their frequencies or orientations. In this article we explain the superposition moiré phenomenon using a spectral model which is based on Fourier analysis. After examining the basic case of cosinusoidal grating superpositions we advance, step by step, through the cases of binary gratings, square grids and dot screens, and discuss the implications on moirés between halftone screens in colour separation. Then, based on these results, we focus on the moiré phenomenon from a different angle, the dynamic point of view: we introduce the moiré parameter space, and show how changes in the parameters of the superposed layers vary the moiré patterns in the superposition. This leads us to an algorithm for moiré minimization which provides stable moiré-free screen combinations for colour separation.

Keywords: Moiré effect, moiré minimization, colour printing, halftoning, Fourier analysis.

1. Introduction

The moiré phenomenon is an optical effect which appears in some circumstances when periodic or quasiperiodic structures (such as line gratings, dot screens, etc.) are intersected. It consists of a visible pattern which is clearly observed at the intersection, although it does not appear in any of the original structures [1]. While this phenomenon has useful applications in several fields, such as in strain analysis or in the detection and measurement of slight deflections or deformations [2, 3, 4], in other situations moiré patterns may have an unwanted, adverse effect. Such is the case in the field of colour separation: moiré patterns which appear in some circumstances between the dot screens used for colour printing may severely deteriorate the quality of the resulting image. It is therefore important to understand the nature and the causes of these moiré patterns in order to know how to avoid, or at least minimize, their adverse effect on colour printing.

Unwanted moiré patterns may appear in the printing process for several possible reasons. For example, if the original image itself is already halftoned, or contains some other periodic fine details, then a moiré pattern may be caused as an interference between the periodic fine details of the original input image and either the input scanning frequency or the printing halftoning frequency. This type of moiré may already appear in black and white printing; but in the field of colour printing, by far the most (in)famous moiré problem is due to the superposition of the halftone screens of the different process colours. This last type of moiré phenomenon is the subject of the present article.

The moiré effect between intersected structures occurs because of the geometric distribution of dark and bright areas in the superposed image: areas in which the dark elements of the original structures cross each other contain less colorant than areas where the original structure elements fall between each other and fill the white spaces better.

Several mathematical approaches have been used to explore the moiré phenomenon. The geometric model [4, 5, 6], which is based on the geometric properties of the superposed layers (their periods and angles) leads to formulas that can predict, under certain limitations, the geometric properties of the moiré patterns. Another approach is the indicial representation of sets of curves [7]; this is a pure algebraic approach, which also yields the same formulas.

For a more founded treatment of the moiré phenomenon, however, an alternative model is required, based on spectral analysis in the *frequency* domain rather than on a geometric or algebraic analysis in the image domain. This spectral model has several

advantages over the simple period-based geometric or algebraic approaches: (1) it lends itself easily to a simple, illustrative graphic representation (the *vector diagram*); (2) using the Fourier transform theory it introduces the possibility of exploring the bidimensional spectrum of the superposed image from which not only the geometric properties of the moiré pattern (angle and period) can be found, as in the other approaches, but also the wave form (*profile*) of the moiré and its strength (*perceptual contrast*); (3) this approach affords a better understanding of the more complex moiré cases and provides an easy explanation of multiple grating moirés, where the geometric analysis may become too complicated. All the above points and terms will be explained and illustrated below.

In this article we will deal basically with the moiré cases which lead to superposition moirés in colour separation. The explanation of this phenomenon is presented here step by step, in a systematic way: we start with the simplest case, the intersection of cosinusoidal gratings, and gradually proceed through the cases of binary gratings and square grids to the intersection of dot screens — which is the basis for the moiré phenomenon in colour separation. This is done based on a useful notational system that we introduce for the identification, classification and labelling of the moiré effects; this formalism is also used for the definition of fundamental terms such as the order of a moiré, singular moiré states, etc. Then, based on these results, we focus in the second part of the article on the moiré phenomenon from a different angle — the dynamic point of view: we introduce the moiré parameter space, and show how changes in the parameters of the superposed layers vary the moiré patterns in the superposition. This leads us to an algorithm for moiré minimization which provides stable moiré-free screen combinations for colour separation. We present the algorithm, and conclude by a discussion of the results obtained.

2. The spectral model – background and basic notions

The spectral model is based on the duality between the original image in the image domain and its spectrum in the frequency domain, through the Fourier transformation. By analyzing properties both in the original image and in its spectral representation this approach can benefit from the advantages of both domains.

In this article we are mainly interested in the bidimensional (2D) case, i.e. in 2D images in the (x,y) plane and their 2D spectra in the (u,v) plane, which are obtained by the 2D Fourier transform. The following paragraphs list the basic properties of the image types we are concerned with and review the implications of these properties both in the image and in the spectral domains.

First, it should be noted that we will only deal here with moiré effects in the monochromatic case; moirés in colour separation will be simulated by the superposition

of the 3 or 4 black-and-white separation films, without taking the ink colours into consideration. In the monochromatic case each printed 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. 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 N monochromatic images are superposed, the reflectance of the resulting image is given by the *product* of the reflectance functions of the individual images:

$$r(x,y) = r_1(x,y) r_2(x,y) \dots r_N(x,y) \quad (1)$$

According to the convolution theorem [9, p. 244; 8, p. 18] the Fourier transform of the product function is the convolution of the Fourier transforms of the individual functions. Therefore if we denote the Fourier transform of each function by the respective capital letter and the 2D convolution by **, we get:

$$R(u,v) = R_1(u,v) ** R_2(u,v) ** \dots ** R_N(u,v) \quad (2)$$

Second, here we are basically interested in *periodic* images, such as gratings or screens, and their superpositions. This implies that the image spectrum is not a continuous one but rather consists of discrete impulses, corresponding to the frequencies which appear in the Fourier series decomposition of the image [9, p. 204]. A strong impulse in the spectrum indicates a pronounced periodic component in the original image at the frequency and direction of that impulse.

Each impulse in the 2D spectrum is characterized by two properties: its *geometric location* (or *impulse location*), and its *amplitude* (see Fig. 1). To the geometric location of any impulse is attached a *frequency vector* \mathbf{f} in the spectrum plane, which connects the spectrum origin with the geometric location of the impulse. This vector can be expressed either by its polar coordinates (f,α) , where α is the direction of the impulse and f is its distance from the origin (i.e. its frequency in that direction); or by its Cartesian coordinates (f_u, f_v) , where f_u and f_v are the horizontal and vertical components of the frequency. In terms of the original image, the *geometric location* of an impulse in the spectrum determines the frequency f and the direction α of the corresponding periodic component in the image, and the *amplitude* of the impulse represents the intensity of that periodic component in the image.

In addition to the periodicity of the original images we assume, unless otherwise mentioned, that each of the original images is *regular* in the sense that its components (lines or dots) are identical, parallel and equidistant.

The periodic images with which we are dealing here are basically of a symmetric nature (gratings, screens, etc.). For the sake of simplicity, we also assume that the given images are not shifted, but indeed centered symmetrically about the origin. As a result, we will deal with images (and image superpositions) which are *real* and *symmetric*, and whose spectra are consequently also real and symmetric [9, pp. 14-15]. This means that each impulse in the spectrum (except for the DC at the origin) is always accompanied by a twin impulse of an identical amplitude, which is symmetrically located at the other side of the origin as in Fig. 1 (their frequency vectors are \mathbf{f} and $-\mathbf{f}$). The complex spectra for the case of non-centered images can be readily obtained whenever required by applying the shift theorem [9, pp. 104-107, 244].

Another important property of our images comes from the fact that most printing devices are only bilevel, i.e. they are only capable of printing solid ink or leaving the paper unprinted, but they cannot produce intermediate ink tones. (This is also true for most colour printing devices, where each of the primary colours is bilevel). In such devices the *visual impression* of intermediate tone levels is usually obtained by means of the halftoning technique, i.e. by breaking the continuous-tone image into small dots whose size depends on the tone level. Therefore, in most practical cases the reflectance function of a printed image is *binary*, only taking values 0 and 1 (signifying the existence or absence of ink on the white paper, respectively).¹ For colour separation we will basically be interested in images consisting of binary dot-screens and their combinations; but before we get there we will first study the basic case of cosinusoidal continuous-tone gratings, and then we will proceed through binary gratings and their intersections to the case of binary dot screens.

3. Superposition of 2 cosinusoidal gratings

Let us first look at the case of gratings with a cosinusoidal wave form. Since reflectance functions always take values between 0 and 1, the cosinusoidal reflectance function has the form (see Fig. 2(a)):

$$r_1(x,y) = 0.5 \cos (2\pi f_1 x) + 0.5 \quad (3)$$

This periodic function has a frequency of f_1 cycles per unit, i.e. its period is $T_1=1/f_1$ units (in the x direction). Similarly, the reflectance function of a cosinusoidal grating with a frequency of f_2 which is rotated by an angle α (Fig. 2(b)) is given by:

¹ It should be noted that in reality, looking under a microscope, the border between black and white areas in an image does not appear to be a sharp transition, but rather a gradual transition in the distribution of the microscopic ink particles. This microscopic view should not, however, invalidate the binary property assumed in the macroscopic view of the image.

$$r_2(x,y) = 0.5 \cos (2\pi f_2[x\cos\alpha + y\sin\alpha]) + 0.5 \quad (4)$$

The 2D Fourier transform of each of these reflectance functions consists of exactly 3 impulses (Fig. 2(d), (e)). In fact, it is the sum of the Fourier transform of the cosinusoidal term which consists of 2 symmetric impulses of amplitude 1/4, located at a distance of $f_i=1/T_i$ from the origin at the respective angle; plus the Fourier transform of the additional constant 0.5, which is an impulse of amplitude 1/2 at the origin (the DC impulse).

According to Eqs. (1) and (2), the spectrum of the superposition of $r_1(x,y)$ and $r_2(x,y)$ is the Fourier transform of their product, i.e. the convolution of their individual spectra $R_1(u,v)$ and $R_2(u,v)$. Since in our cosinusoidal case each of these spectra consists of 3 impulses, their convolution consists of 9 impulses (see Fig. 2(f)). Using the graphical method of “move and multiply” [9, pp. 29-30; 8, pp. 13-14] we see that the convolution is 0 throughout the (u,v) plane, except at the points where the impulses of the moving copy of $R_2(u,v)$ fall whenever the origin of $R_2(u,v)$ is placed on top of an impulse of $R_1(u,v)$. This means that the *geometric location* of the impulses in the convolution can be found simply by placing on top of each impulse of $R_1(u,v)$ a centered copy of $R_2(u,v)$ (or vice versa, since convolution is commutative). The *amplitude* of each impulse thus received is the product of the amplitudes of the 2 impulses involved: the impulse in the first spectrum on top of which the moving spectrum is centered, and the impulse in the moving spectrum which then defines the location of the impulse in question. If a newly generated impulse falls on top of an already existing impulse, their amplitudes are summed up. The amplitudes received in our case are 1/4 for the DC impulse, 1/8 for the 2 impulse pairs of the original cosines, and 1/16 for the 2 new impulse pairs generated by the convolution (see bottom row in Fig. 2).

The spectrum resulting from the convolution contains all the impulse pairs of the original spectra (only their amplitudes have been modified, but not their geometric locations). However, 2 new pairs of impulses which did not exist in any of the original spectra have appeared in the convolution (see Fig. 2(f)). The geometric locations of these new impulse pairs are determined by the vectorial sum and the vectorial difference of the frequency vectors of the original impulse pairs, namely: $\mathbf{f}_1+\mathbf{f}_2$, $-\mathbf{f}_1-\mathbf{f}_2$, and $\mathbf{f}_1-\mathbf{f}_2$, $\mathbf{f}_2-\mathbf{f}_1$.

Since each impulse pair in the spectrum reflects a periodic component with the corresponding frequency and angle in the original image, these 2 new impulse pairs suggest that the superposition of the 2 original images includes 2 new periodic components which did not exist in either of the original images. And indeed, looking at the superposed image (Fig. 2(c)) we can identify these new periodic components: The more obvious one has the frequency and direction of the difference vector, $\mathbf{f}_1-\mathbf{f}_2$; but the other one, with the frequency and the direction of the sum vector, $\mathbf{f}_1+\mathbf{f}_2$, can also be identified. The first periodic component is more visible than the other since its

frequency is lower, i.e. its period is larger. While the frequency of the vector sum is always larger than the frequencies of each of the individual vectors, the frequency of the vector difference may be significantly smaller than either of the original frequencies. Consequently, the periodic component in the superposed image which corresponds to the vector difference can have a significantly larger period, and therefore be much more visible, than the cosines of the original images. This prominent periodic component is, in fact, the moiré effect seen in the superposition of the 2 original images (Fig. 2(c)).

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° [10, p. 531]). However, for the sake of simplicity this low-pass filter can be approximated by the *visibility circle*, a circular step function 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 exact radius depends on several factors such as the contrast of the observed details, the viewing distance, light conditions, etc. 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 the frequency of the vectorial difference falls inside the visibility circle, then a moiré effect appears in the superposed image. (In fact, the visibility circle has a hole in its center, since very low frequencies cannot be seen, either).

4. Superposition of three or more cosinusoidal gratings; singular moiré states

If we now superpose a third cosinusoidal grating on top of the first two, the resulting spectrum will be the convolution of all three spectra — i.e. the result of convolving the third spectrum with the convolution of the first two spectra. The geometric location and the amplitude of the impulses in the resulting spectrum can be found graphically in the same manner as above: a centered copy of the new 3-impulse spectrum of the third cosinusoidal grating is placed on top of each of the 9 impulses of the 2-layer convolution, thus generating 9 additional impulse pairs in the combined spectrum (Fig. 3). The amplitude of each of the impulses of the copied spectrum is scaled by the amplitude of the impulse on top of which it has been copied. If any of the newly generated impulses fall inside the visibility circle, a new periodic component (or moiré effect) can be seen in the image.

Generalizing this to the superposition of N cosinusoidal gratings, we see that the final convolution contains the frequency vectors of each of the original images as well as all the frequency vectors obtained in each successive convolution. This means that the final

convolution contains all the frequency vectors which can be obtained as a vectorial sum of 1, 2, ... or N frequency vectors, one (or none) from each original spectrum. If we consider the DC impulse of each spectrum as having a zero frequency vector, it can be said that *each* of the individual spectra contributes *one* of its frequency vectors to every vectorial sum. In other words, the frequency vector \mathbf{f} of any individual impulse in the final convolution (i.e. in the spectrum of the superposition of the N images) is a vectorial sum of N frequency vectors \mathbf{f}'_i , where \mathbf{f}'_i is one of the 3 frequency vectors contained in the spectrum of the i -th image:

$$\mathbf{f} = \mathbf{f}'_1 + \mathbf{f}'_2 + \dots + \mathbf{f}'_N \quad (5)$$

If θ_i are the angles that the N original cosinusoidal images form with the positive horizontal axis and f_i are their frequencies, then the two Cartesian components of the above vectorial sum can be written as:

$$\begin{aligned} f_u &= f_1 \cos\theta_1 + f_2 \cos\theta_2 + \dots + f_N \cos\theta_N \\ f_v &= f_1 \sin\theta_1 + f_2 \sin\theta_2 + \dots + f_N \sin\theta_N \end{aligned}$$

and the frequency, the period and the angle of the impulse in question are given by the length and the direction of the sum vector \mathbf{f} :

$$f = \sqrt{f_u^2 + f_v^2} \quad T = 1/f\phi = \arctan(f_v/f_u) \quad (6)$$

The way in which any frequency vector \mathbf{f} in the spectrum of the superposition is obtained from frequency vectors \mathbf{f}_i in the individual spectra of the original images can be shown graphically using the geometric rules of vector addition, as shown in Fig. 10. Such *vector diagrams* can be drawn for any frequency vector \mathbf{f} in the spectrum of the superposed image, but they are of particular interest for those impulses in the spectrum which correspond to moirés in the superposed image. The vector diagrams provide a clear illustrative explanation of the nature of any moiré in question.

If the impulse whose frequency vector is \mathbf{f} falls inside the visibility circle and represents a visible moiré in the superposition of the N original images, the above formulas (6) in fact express the frequency, the period and the angle of this moiré. Note that in the special case of $N=2$, where a moiré effect occurs due to the vectorial sum of the frequency vectors \mathbf{f}_1 and $-\mathbf{f}_2$, these formulas are reduced to the familiar geometrically obtained formula of the moiré effect between 2 layers [4]:

$$T = \frac{T_1 T_2}{\sqrt{T_1^2 + T_2^2 - 2T_1 T_2 \cos\alpha}} \quad (6.1)$$

(where T_1 and T_2 are the periods of the 2 original images and α is the angle difference between them, $\theta_1 - \theta_2$).

Eqs. (5) and (6) only describe the *geometric properties* of an impulse in the spectrum of the superposition (and of the periodic component or moiré which it represents in the image domain). The *amplitude* of any individual impulse, which represents the strength of the corresponding periodic component in the image, is a product of the amplitudes of the N impulses from which it has been received in the convolution, one from each of the N spectra:

$$a = a_1 a_2 \dots a_N \quad (7)$$

Note however that if 2 or more impulses in the convolution happen to fall on top of each other in the same location, their individual amplitudes are summed up.

As can be seen from Eqs. (5) and (7), the spectrum convolution can be seen as an operation in which frequency vectors of the original spectra are added vectorially, whereas the corresponding impulse amplitudes are multiplied. Note that since all the convolved spectra are real and symmetric about the origin (see Sec. 2), the resulting spectrum is also real and symmetric, and contains for each impulse at location \mathbf{f} an identical twin impulse at $-\mathbf{f}$.

We have seen that if one or several of the new impulse pairs in the convolution fall close to the origin, inside the visibility circle, they imply the existence of one or several moirés with significant periods in the superposed image (Fig. 3(b), (d)). An interesting special case occurs when some of the impulses of the convolution fall *exactly* on top of the DC impulse at the origin. This happens for instance in the superposition of 2 identical gratings at the same angle, or when 3 identical gratings are superposed at angle differences of 120° between each other. As can be seen from the vector diagrams, these are limit cases in which the vectorial sum of the frequency vectors is exactly 0. This means that the moiré frequency is 0 (i.e. its period is infinitely large), and therefore the moiré is not visible. This situation is called a *singular moiré state*; but although the moiré effect in a singular state is not visible, this is a very unstable moiré-free state since any slight deviation in the angle or frequency of any of the superposed images may cause the new impulses in the spectrum to move slightly from the origin, thus generating a moiré effect with a very significant period [6, Figs. 6,8].

5. Binary square waves and their spectra

Let $r(x)$ be a one dimensional binary (0, 1 valued) periodic square wave. We will denote the period of this function by T and the width of its *white* square pulse (the *opening*) by τ (see Fig. 4). As usual, we assume that the square wave is symmetrically centered around the origin, so that both the original image and its spectrum are *real* and *symmetric*.

A periodic square wave can be expressed by means of its Fourier series expansion, as an infinite series of weighted sine and cosine functions, at the fundamental frequency of $1/T$ and all its harmonics. The general expansion (or decomposition) of a one dimensional periodic function into a Fourier series is given by:

$$r(x) = a_0 + 2 \sum_{n=1}^{\infty} a_n \cos(2\pi nx/T) + 2 \sum_{n=1}^{\infty} b_n \sin(2\pi nx/T) \quad (8)$$

[9, p. 205], where the Fourier series coefficients are given by:

$$a_n = \frac{1}{T} \int_0^T r(x) \cos(2\pi nx/T) dx \quad b_n = \frac{1}{T} \int_0^T r(x) \sin(2\pi nx/T) dx$$

(the integration may be done over any 1-period interval, such as $-T/2 \dots T/2$, etc). In the special case of a symmetric binary square wave, where $r(x)=1$ between $-\tau/2 \dots \tau/2$ and 0 elsewhere (see Fig. 4), these coefficients are:

$$(9) \quad \begin{aligned} a_0 &= \tau/T \\ a_n &= (1/n\pi) \sin(n\pi\tau/T) = (\tau/T) \text{sinc}(n\tau/T) \\ b_n &= 0 \end{aligned}$$

(there are no *sine* components here, due to the symmetry of the square wave). The fact that the square wave can be expressed as a constant a_0 plus an infinite sum of cosine functions, implies that the Fourier transform of the square wave contains a DC impulse whose amplitude is a_0 , plus an infinite series of impulse pairs located at the frequencies of $\pm n/T$. Since the n -th impulse pair is the Fourier transform of the n -th cosine in the Fourier series decomposition of the square wave (the n -th harmonic), its amplitude is given by the n -th cosine coefficient, a_n (Eq. 9). The spectrum of the square wave is therefore:

$$R(u) = \sum_{n=-\infty}^{\infty} a_n \delta(u-n/T) \quad (10)$$

(where $\delta(u)$ is the impulse symbol). This is an impulse train (or a ‘‘Dirac comb’’) which samples the continuous function (‘‘envelope’’) $g(u)=(\tau/T)\text{sinc}(\tau u)$ at the fundamental frequency $u=1/T$ and all its harmonics, $u=n/T$ (Fig. 4). The amplitude of the impulses oscillates and fades out symmetrically in both directions from the center.

Note that the period T of the square wave determines the interval $1/T$ between each two successive impulses of the comb in the spectrum, while the opening τ ($0 \leq \tau \leq T$) determines the length of the lobes in the envelope sinc function. The height of the envelope at the origin, i.e. the amplitude of the DC impulse, is determined by the *opening ratio* of the square wave, τ/T ($0 \leq \tau/T \leq 1$).

An interesting observation from the formula of a_n (Eq. 9) is that when the opening ratio τ/T of a given square wave is rational, i.e. if it can be expressed as l/k where l and k are integers, then for any n which is a multiple of k the impulse amplitude a_n is zero.

For instance, if $\tau/T=1/2$ then every even impulse in the comb has a 0 amplitude, and if $\tau/T=1/4$ or $3/4$ then every fourth impulse in the comb is zero.

Let us mention here one more result that will be used later: If $R(u)$ is the Fourier transform of a symmetric square wave $r(x)$, then the Fourier transform of the “negative” wave $1-r(x)$ is given by:

$$F(1-r(x)) = F(1) - F(r(x)) = \delta(u) - R(u) = \begin{cases} 1-R(0) & u=0 \\ -R(u) & u \neq 0 \end{cases} \quad (11)$$

Therefore, if a_n are the impulse amplitudes in the spectrum of $r(x)$, then the impulse amplitudes in the spectrum of $1-r(x)$ are: $c_0 = 1-a_0$, $c_n = -a_n$.

Proceeding now to the 2D case, the general expansion of a periodic function $r(x,y)$ with periods T_x and T_y in the x and y directions into a 2D Fourier series is given by:

$$r(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_{m,n} \cos 2\pi(mx/T_x + ny/T_y) + \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} b_{m,n} \sin 2\pi(mx/T_x + ny/T_y)$$

$$\text{where: } a_{m,n} = \frac{1}{T_x T_y} \int_0^{T_x} \int_0^{T_y} r(x,y) \cos 2\pi(mx/T_x + ny/T_y) dx dy$$

$$b_{m,n} = \frac{1}{T_x T_y} \int_0^{T_x} \int_0^{T_y} r(x,y) \sin 2\pi(mx/T_x + ny/T_y) dx dy$$

If $r(x,y)$ is symmetric, the infinite sum only contains cosine terms, each of which is a function of both mx and ny . Each of these cosine terms, multiplied by a coefficient $a_{m,n}$, gives (isolates) the contribution to the image $r(x,y)$ of the cosinusoidal periodic component in the direction and frequency of the (m,n) -th harmonic. The spectrum (Fourier transform) of $r(x,y)$ is, therefore, an impulse nailed which contains for each of the (m,n) -th harmonics an impulse of amplitude $a_{m,n}$. The frequency vector of each impulse in the nailed denotes the direction and the frequency of the corresponding periodic component in the image. In the following section we will discuss the special case where $r(x,y)$ is a binary square wave. Note that since a 2D square wave is basically a constant extension of a 1D square wave into the perpendicular direction, its 2D Fourier series representation (and therefore its 2D spectrum) are in fact of 1D nature.

6. Superposition of binary gratings; higher order moirés; the notational system

The reflectance function of a binary grating is a bidimensional binary periodic square wave; for the sake of convenience we will assume it is also symmetric. The reflectance function of a binary grating with period T_1 and opening τ_1 in direction x (Fig. 5(a)) is given by the Fourier series:

$$r_1(x,y) = a^{(1)}_0 + 2 \sum_{n=1}^{\infty} a^{(1)}_n \cos(2\pi n x/T_1) \quad (12)$$

where the coefficients $a^{(1)}_0$ and $a^{(1)}_n$ are similar to a_0 and a_n in Eq. (9) above, with only T_1 and τ_1 replacing T and τ . Similarly, the reflectance function of a binary grating with period T_2 and opening τ_2 , which is rotated by angle α (Fig. 5(b)), is given by the Fourier series:

$$r_2(x,y) = a^{(2)}_0 + 2 \sum_{n=1}^{\infty} a^{(2)}_n \cos(2\pi n [x \cos \alpha + y \sin \alpha]/T_2) \quad (13)$$

where $a^{(2)}_0$ and $a^{(2)}_n$ are similar coefficients with T_2 and τ_2 .

The Fourier transform $R_1(u,v)$ of the reflectance function $r_1(x,y)$ is a symmetric 1D impulse train on the u axis (Fig. 5(d)), whose envelope is $g(u) = (\tau_1/T_1) \text{sinc}(\tau_1 u)$. The intervals between the impulses are $1/T_1$ and their amplitudes are $a^{(1)}_n$. Similarly, the Fourier transform $R_2(u,v)$ of the reflectance function $r_2(x,y)$ is a 1D symmetric impulse train with an envelope shape of $g(u) = (\tau_2/T_2) \text{sinc}(\tau_2 u)$ on a line in the u,v plane which has been rotated by angle α , with an impulse interval of $1/T_2$ and an impulse amplitude of $a^{(2)}_n$ (Fig. 5(e)).

Let us now consider the superposition of the 2 binary gratings, $r(x,y) = r_1(x,y)r_2(x,y)$. According to the convolution theorem (Eq. 2) its spectrum $R(u,v)$ is the convolution of spectra $R_1(u,v)$ and $R_2(u,v)$. This convolution can be carried out graphically by the “move and multiply” method as in the case of 2 cosinusoidal functions, the only difference being that now each of the individual spectra contains an infinite number of impulses (a “comb”) rather than only 3 impulses. The result of the convolution, in terms of impulse locations, is an infinite oblique lattice of points (frequency vectors) which results from placing a centered, parallel copy of the comb $R_2(u,v)$ on top of each impulse of the comb $R_1(u,v)$, or vice versa (see Fig. 5(f), and compare with the cosinusoidal case shown in Fig. 2(f)). More precisely, the results of the convolution can be formulated as follows:

- (a) The *frequency vectors* of the impulses in the spectrum of the superposed gratings are the vectorial sums of each frequency vector on the comb of the first grating and each frequency vector on the comb of the other grating. All these frequency vectors (or rather their corresponding impulse locations) form the above-mentioned lattice.
- (b) The *amplitude* of the (m,n) -th impulse on this oblique lattice is the product of the amplitudes of the m -th impulse in the first comb and the n -th impulse in the second comb:

$$\begin{aligned}
 a_{m,n} &= a_m^{(1)} a_n^{(2)} \\
 &= (1/\pi^2 mn) \sin(\pi m \tau_1/T_1) \sin(\pi n \tau_2/T_2) \\
 (14) \quad &= (\tau_1 \tau_2/T_1 T_2) \text{sinc}(m \tau_1/T_1) \text{sinc}(n \tau_2/T_2)
 \end{aligned}$$

As in the cosinusoidal case, we see that here too the superposition of gratings introduces new impulses in the spectrum. If any of these impulses fall inside the visibility circle, as in Fig. 5(f), it indicates that in the superposed image there is a visible periodic component (i.e. a moiré effect) at the corresponding direction and frequency.

Results (a) and (b) above lead to the following interesting observations:

(1) The angle difference between the two gratings only influences the geometric location of the impulses, but not their amplitudes. In the image domain this means that the angle difference between the two gratings only influences the angle and period of each moiré, but not its amplitude. (In case of more than 2 superposed gratings, however, angle changes may cause some impulses to fall on top of each other, in which case their amplitudes are summed up).

(2) Varying the opening τ of any of the gratings while the period T remains fixed only influences the amplitude of the impulses, but not their location. In the image domain this means that only the amplitude of the moiré is influenced, but not its angle or period.

The generalization of the results obtained above to the superposition of N binary gratings is straightforward. In fact, the geometric location of each impulse in the resulting spectrum is determined by the vectorial sum of frequency vectors, one from each of the superposed gratings, while the amplitude of the individual impulses is the product of the amplitudes of the involved impulses. This is very similar to the case of cosinusoidal gratings, except that the spectrum of a binary grating consists of an infinite comb of impulses rather than only 3 impulses, and therefore in the case of binary gratings the convolution of N spectra gives an infinite nailbed of impulses rather than just a finite number (3^N) of impulses. This means that in the case of binary gratings each of the components \mathbf{f}'_i in Eq. (5) may come from any impulse of the comb of the i -th spectrum. In fact, if \mathbf{f}_i is the frequency vector of the fundamental impulse in the i -th spectrum, then the frequency vector of its k_i -th harmonic impulse is $k_i \mathbf{f}_i$, and Eqs. (5) and (7) for the general (k_1, k_2, \dots, k_N) -impulse in the convolution become:

$$\mathbf{f}_{k_1, k_2, \dots, k_N} = k_1 \mathbf{f}_1 + k_2 \mathbf{f}_2 + \dots + k_N \mathbf{f}_N \quad (15)$$

$$\begin{aligned} a_{k_1, k_2, \dots, k_N} &= a_{k_1}^{(1)} a_{k_2}^{(2)} \dots a_{k_N}^{(N)} \\ &= \frac{1}{\pi^N k_1 k_2 \dots k_N} \sin(\pi k_1 \tau_1 / T_1) \dots \sin(\pi k_N \tau_N / T_N) \\ &= \frac{\tau_1 \tau_2 \dots \tau_N}{T_1 T_2 \dots T_N} \operatorname{sinc}(k_1 \tau_1 / T_1) \dots \operatorname{sinc}(k_N \tau_N / T_N) \end{aligned} \quad (16)$$

where k_i are integer numbers. If 2 or more impulses in the convolution happen to fall on the same point (i.e. they have identical frequency vectors), their individual amplitudes are summed up.

The Cartesian components of the vectorial sum in (15) are:

$$\begin{aligned} f_{U, k_1, k_2, \dots, k_N} &= k_1 f_1 \cos \theta_1 + k_2 f_2 \cos \theta_2 + \dots + k_N f_N \cos \theta_N \\ f_{V, k_1, k_2, \dots, k_N} &= k_1 f_1 \sin \theta_1 + k_2 f_2 \sin \theta_2 + \dots + k_N f_N \sin \theta_N \end{aligned} \quad (17)$$

and as in the cosinusoidal case they can be inserted into Eq. (6) in order to obtain the frequency, the period and the angle of the impulse in question (and of the moiré it may represent, in case it falls inside the visibility circle).

As a result, two main differences emerge between the superpositions of cosinusoidal or binary gratings: first, in the case of binary gratings, each impulse in the visibility circle (like any other impulse in the spectrum) belongs to an infinite comb of impulses, in the same direction, through the origin (see Fig. 5(f)). This means that each moiré is represented in the spectrum by an *infinite* series of impulses. The fundamental impulse (i.e. the first impulse next to the DC) in this comb determines the period and the direction of the moiré, and therefore it is still sometimes called, as in the cosinusoidal case, “the impulse of the moiré”. If further harmonic impulses of this comb also fall inside the visibility circle, the profile shape of the moiré is no longer perceived as a pure cosinusoidal function, but rather as a more complex form (a sum of cosines).

Second, in the case of binary grating superposition, the visibility circle may also contain impulses which originate from higher harmonic impulses in the combs of the individual spectra. This means that, unlike in the cosinusoidal case, moiré effects between binary gratings can also be obtained from higher harmonics of the fundamental grating frequencies. Such moiré effects are called *higher order moirés* [11]. This is demonstrated for the case of 2 superposed gratings in Fig. 6; note that the visible moiré effect is caused here by the vectorial sum (or rather difference) of \mathbf{f}_1 , the fundamental frequency of the first grating, and *twice* \mathbf{f}_2 , i.e. the second harmonic of the other grating, while the vectorial difference $\mathbf{f}_1 - \mathbf{f}_2$ is outside the visibility circle. In the image domain

this means that the visible moiré is actually due to the intersection of every *second* line in the first grating with each line of the second grating. If both $\mathbf{f}_1 - \mathbf{f}_2$ and $\mathbf{f}_1 - 2\mathbf{f}_2$ were inside the visibility circle, then both the first and the second order moirés could be observed simultaneously in Fig. 6(c). Fig. 7 shows some of the various order moirés which may exist between 2 superposed gratings.

It will be useful to present at this point the systematic notational formalism that we introduce for the following discussions, which provides an unambiguous means for the identification, classification and labelling of the moiré effects. We will call the N -grating moiré whose fundamental impulse is the (k_1, k_2, \dots, k_N) -impulse in the spectrum convolution (see above) a (k_1, k_2, \dots, k_N) -moiré; its comb contains the convolution impulses $(nk_1, nk_2, \dots, nk_N)$ for any integer n (this can be seen for $N=2$ in Figs. 5, 6). This moiré is singular if $\sum k_i \mathbf{f}_i = \mathbf{0}$. Note that sometimes, when no confusion may occur, it is convenient to use a shorthand notation in which zero indices and negative signs are omitted. For instance, the $(2,0,-1,1)$ -moiré between 4 gratings may be simply called a $\{2,1,1\}$ -moiré, and both the 4 grating $(1,0,-1,0)$ -moiré and the 2 grating $(1,-1)$ -moiré may be called in short $\{1,1\}$ -moirés. This shorthand notation is sometimes more convenient than enumerating all the possible similar variants. The highest absolute value in the list is called the *order* of the moiré.

7. The profile of the moiré and its perceptual contrast

As has been shown above, each moiré in the superposition of binary gratings is represented in the spectrum by an *infinite* series of impulses (a Dirac comb). The amplitudes of these impulses are the coefficients of the Fourier series development of the function which represents the *profile* of the moiré (i.e. the intensity of the moiré at any point along a cross-section of the image in the moiré direction). The exact profile shape of the moiré in the image domain can be reconstructed as an infinite sum of cosines, by inserting the amplitudes $a_n = a_{nk_1, nk_2, \dots, nk_N}$ of the moiré comb impulses given by Eq. (16) into Eq. (8) (with $b_n = 0$ for any n , since we are dealing with symmetric cases). An approximate profile shape of the moiré can be obtained from the first few impulse pairs of the comb which fall inside the visibility circle, i.e. as a sum in the image domain of only the first few cosine terms in the Fourier series (8) (plus a constant term due to the DC impulse, which represents the average level of the profile). A coarser, sinusoidal shaped approximation of the moiré profile can already be obtained from the DC and the fundamental impulse pair alone (whose frequency vector gives the basic angle and frequency of the moiré).

As we have seen, our mathematical model assigns to each point of the image a reflectance value between 0 and 1, where 0 means black, 1 means white, and intermediate values represent in-between reflectance values. This applies, in particular, to the moiré profiles and the image projections which are calculated as described above

based on this model.² Note that although the original images and their superpositions are all binary and only take values 0 and 1, their moiré profiles may also have intermediate values. In fact, the value of each point on a profile represents the average ratio of white per unit of length (or area), i.e. the average reflectance at that point. For example, in the (1,-1)-moiré between 2 identical gratings where the width of the black and white lines is identical (i.e. the opening ratio is $\tau/T=0.5$; see Fig. 8(b)), then the value of the moiré profile along the center of the dark moiré bands is 0 (no white at all), and this value gradually climbs up to 0.5 at the center of the bright moiré bands (where black and white are equally distributed, and therefore the white ratio is 0.5). This is shown graphically in the reflectance profile of the moiré, in Fig. 8(e); similar profiles are shown in Figs. 8(d),(f) for other opening ratios corresponding to the moirés of Figs. 8(a),(c).

However, the difference between the maximum and the minimum reflectance of the moiré profile does not represent correctly the degree of contrast as it is perceived by the human eye. For example, in the (1,-1)-moiré between 2 identical gratings, the difference between the maximum and the minimum reflectance profile values are identical for gratings with opening ratios of $\tau/T=0.75$ and $\tau/T=0.25$, while the eye clearly sees a much higher contrast in the second (darker) case (compare Figs. 8(a),(c) with their reflectance profiles 8(d),(f)).³ The reason for this phenomenon is that the response (or sensibility) of the human visual system to light intensity is not linear in its nature, but rather close to logarithmic [12, pp. 27-29]. If we plot the intensities or the moiré profiles logarithmically, i.e. in terms of density rather than in terms of reflectance, we get a more realistic representation of the perceptual contrast of the moiré which better corresponds to the human perception (see Fig. 8(g)-(i)). A still better correspondence can be achieved by replacing the logarithmic approximation of the human visual response by an empiric function based on the experimental data obtained from physiological research [13, pp. 62-69].

8. Square grids and their superpositions

A square grid is in fact an orthogonal superposition (multiplication) of two identical binary gratings, which together form a pattern of identical squares. Therefore the spectrum of a square grid is the convolution of two identical but perpendicular combs, and its impulses are located on a square lattice. The amplitude of each impulse in the

² The range of [0,1] is respected, of course, only by the *precise* profile functions which take into account all the impulses till infinity. An approximation using only a finite number of terms, such as the DC plus the first harmonic cosine (x2), may somewhat exceed the range of [0,1].

³ Note that even the Michelson contrast [13 p. 34], defined as: $(\max-\min)/(\max+\min)$, which is widely used in literature, is not appropriate here — since for both cases 8(e),(f) it yields the same value, 1.

spectrum is given by Eq. (14); as we can see, the spectrum has the form of a nailbed which samples the continuous bidimensional function (“envelope”) $g(u,v) = (\tau/T)^2 \text{sinc}(\tau u) \text{sinc}(\tau v)$ at the points $(u,v) = (m/T, n/T)$, where T and τ are the period and the opening of both gratings. The shape of such a spectrum can be seen in Fig. 11(f).

The spectrum of a superposition of two square grids is the convolution of 2 such nailbeds: a centered copy of one of the nailbeds is scaled and placed on top of each impulse of the other nailbed. But the superposition of 2 square grids can also be seen as a special case of 4 binary grating superposition with two pairs of gratings, each pair consisting of 2 identical but perpendicular gratings. It is clear, therefore, that each impulse pair in the spectrum of the superposition is accompanied by an identical, perpendicular impulse pair. This means that every moiré effect which occurs between any of the gratings involved always has a perpendicular twin; the visual effect in the superposed image therefore has a 90° symmetry (such as a square grid or screen).

The moiré effects which occur in the superposition of two square grids may originate from 2, 3, or from all of the 4 gratings involved. This means that the frequency vector of the fundamental moiré impulse in the visibility circle of the spectrum may be a vectorial sum of 2, 3, or 4 frequency vectors.⁴ Examples of these types of moirés between 2 square grids are shown in Fig. 9; their vector diagrams are shown in Fig. 10. In each case, the low frequency vectorial sum of the indicated frequency vectors from the 2 original spectra gives the fundamental frequency vector of the visible moiré effect, and therefore determines the angle and the frequency of the moiré. (Note that in order to determine the exact profile shape of the moiré, its higher harmonic impulses are also required, insofar as they are situated inside the visibility circle).

The generalization into the case of 3 or more superposed square grids is straightforward. The general expressions for the frequency vector and the amplitude of an impulse in the spectrum of the superposition are given by Eqs. (15, 16), where N is the number of gratings involved, i.e. twice the number of the square grids.

9. Dot screens and their superpositions

A *regular dot screen* is a dot screen with perpendicular axes and equidistant, identical dots. This is in fact a generalization of a square grid, since the square grid can be considered as a special case of a screen having square, white dots on black background (see Fig. 11(e)). A regular screen $r(x,y)$ of white dots on a black background can be

⁴ A grating which does not explicitly participate in the generation of the moiré effect in fact contributes its DC impulse: it contributes a zero frequency vector to the vectorial sum of the frequency vectors, and the amplitude of its DC impulse to the product amplitude (see Eqs. 15, 16).

seen as a convolution of a “spot function” $d(x,y)$ which describes a single white dot, with a nailbed of period T (see Fig. 11). Therefore, according to the convolution theorem, the spectrum of such a dot screen is the product of the continuous function $D(u,v)$ (the Fourier transform of the spot function), and the Fourier transform of the nailbed which is itself a scaled and stretched nailbed. This product is a nailbed which samples the “envelope” function $D(u,v)$ at intervals of $1/T$, scaling its amplitude by $1/T^2$ (see Fig. 11). In the case of a square white dot whose side is τ (as in Fig. 11), the spot function is: $d(x,y) = {}^2\Pi(x/\tau,y/\tau)$ (where ${}^2\Pi(x/\tau,y/\tau)$ denotes the function whose value is 1 within the square $-0.5 \leq x,y \leq 0.5$ and 0 elsewhere), and therefore the envelope of the spectrum is: $D(u,v) = \tau^2 \text{sinc}(\tau u) \text{sinc}(\tau v)$ [9, p. 246]. In the case of a circular dot with radius τ , the envelope $D(u,v)$ is given by a Bessel function [9, p. 248], which is sometimes called the *sombrero function* due to its circular symmetry.

As we can see, the shape of an individual dot $d(x,y)$ of the screen determines the shape of the envelope function $D(u,v)$ in the spectrum, and therefore it determines the amplitude of each of the impulses in the spectrum of the screen. However, the geometric locations of the impulses in the spectrum of the screen are not influenced by the dot shape, and they are determined only by the nailbed with which the dot is being convolved.

A regular screen of *black dots on white background*, which represents a halftone screen with dot area coverage of less than 50%, can be seen as the “negative” of the above white-on-black screen $r(x,y)$, namely: $1 - r(x,y)$. Following Eq. (11), the Fourier transform of $1 - r(x,y)$ contains exactly the same impulses as $R(u,v)$, the spectrum of the square grid $r(x,y)$, where only the sign of their amplitudes have been inverted: $c_{m,n} = -a_{m,n}$, and the DC impulse has an amplitude of: $c_{0,0} = 1 - a_{0,0}$. Here again, only the impulse amplitudes have been influenced, but not their geometric locations.

The spectrum of the superposition of 2 dot screens is the convolution of the 2 bell-shaped nailbed spectra of the individual screens, as in the special case of square grids. Such a convolution process in the spectrum is repeated for each additional dot screen which is superposed in the image domain. If after the convolution the spectrum contains new impulses inside the visibility circle which yield sufficiently strong perceptual contrasts, then new visible periodic components with the respective angles and frequencies appear in the superposed image as moiré effects. This is very similar to the special case of square grids discussed above (i.e. the case of square white dots on a black background), and in fact the moiré effects received in both cases have very similar *macroscopic* properties (see Fig. 9). Note however that the *microscopic* properties of the two cases may be quite different. The microscopic structure in the superposition of dot screens is distinguished by a fine pattern of small dot clusters, called *rosettes* [5], which can be seen even when no visible moirés are present.

The fact that the geometric location of the impulses in the spectrum of the screen does not change when the form of the individual dot is modified (or when the negative of the screen is taken) is fundamental for the understanding of the moiré effects between superposed screens. It means that the period and the direction of the moiré do not depend on the shape of the individual dots of the superposed screen; changing the dot shape (including the dot size) may only affect the *intensity* and the *profile* of the moiré effect. This is shown in Fig. 12.

10. The implications on moirés between halftone screens in colour separation

A halftoned image with clustered screen elements is a screen of equidistant dots, in which the dot size is not constant but rather varies (according to the tone values in the original image). The moirés between halftoned images are therefore based on the case of regular dot screens which has already been described above. Since the colour separation technique involves overprinting 3 or 4 halftoned images, one for each of the primary colours [10, Sec. 26.6], undesired moiré patterns may occur between these halftone screens if special care is not taken.

In order to be clearly visible, a moiré between 2 or more screens should have a large enough period (i.e. its fundamental impulse in the spectrum must be inside the visibility circle), and its perceptual contrast (see Sec. 7) should be relatively strong. However, it is important to note that in colour separation not only strong moirés with large periods may be harmful. Weaker moirés with large periods (say, 10 times the screen period or more) can still be visible, as well as strong moirés with small periods (3-5 times the screen period); the latter may cause a rough linen-like texture and give the printed image a grainy aspect. As a rule of thumb it can be said that pronounced moiré structures which are larger than 1 mm are already clearly visible in normal viewing conditions and therefore undesirable.

The minimization of moirés between superposed screens consists of finding parameter combinations for the superposed layers which give as far as possible moiré-free results. But the task of finding a good moiré-free combination of more than 2 halftone screens is not an easy one, since many different moirés may appear at each angle and frequency combination (this is illustrated in Figs. 7 and 13 for the case of 2 superposed layers). Minimizing one of these moirés (by varying the screen frequencies and angles) does not guarantee the minimization of the others, and in fact the task of finding a good screen combination is a trade-off between the contradictory tendencies of the various moirés involved.

It should also be noted that not all the moiré minima are indeed stable solutions. As described in Sec. 4 many of the moiré-free combinations are in fact singular moiré

states; therefore any small deviation in the angle or frequency of any of the dot screens may cause the reappearance of the moiré pattern in its full strength. Singular moiré-free states may be used in colour separation only if high accuracy can be guaranteed for the screen angles and frequencies; in fact, such is the case in the traditional screens with equal frequencies and equal angle distribution (30° - 30° - 30°). But moiré-free combinations which are not singular should be preferable, since they are more stable and much less sensitive to small inaccuracies in the reproduction and printing process.

In halftoned images, different gray levels are obtained by variations in the size and sometimes also in the shape of the screen dots. But as we have seen, the envelope shape and therefore the amplitude of each impulse in the nailed spectrum of a dot screen are strongly influenced by the shape of the dot, as well as by the opening ratio τ/T or the dot size. This means that the intensity and the profile of the moirés between halftone screens (but not their angles or periods!) are influenced by the shape and size of the dots, since they depend on the amplitude of the new impulses received in the superposition of these screens. In other words, even at a fixed frequency and angle screen combination the intensity and the profile of the generated moiré (or moirés) may still vary significantly (within their fixed periods and angles) as the halftone gray levels are being changed. Furthermore, when several moirés are present simultaneously, each of them may be dominant in a different combination of gray levels. These facts are clearly demonstrated in Fig. 13. It is therefore not sufficient to look for good screen combinations only at some predefined gray levels, since an innocent moiré in one gray level may prove to be much stronger in another gray level. However, an exhaustive run through a large range of screen parameters and gray levels in order to find the best screen combinations proves to be impractical: due to the large amount of calculation, and also due to the fact that for most of the practically used halftone cells (such as black circles which gradually grow into squares at 50% and then gradually turn into white circles) we have no precise expressions for the impulse amplitudes. We therefore selected a simpler approach for the moiré minimization, which is only based on geometric parameters such as angles and frequencies (although perceptual intensities of the moirés are also taken implicitly into account by experimental measurements). This approach will be explained in the following sections.

11. Dynamics of the moiré phenomenon; the moiré parameter space

In the foregoing sections we analyzed the moiré phenomenon between superposed layers only from the *static* point of view. That is, given a fixed combination of angles and frequencies for the individual layers we analyzed the effects of their superposition, both in the image domain and in the spectral domain (like in Figs. 5, 6, 9-10).

Based on these results, in the present section we will consider the *dynamic* aspects of the moiré phenomenon between superposed layers, i.e. how changes in the parameters

of the superposed layers affect the moiré phenomena in the superposition, both in the image and in the spectral domains.⁵ Understanding these dynamic aspects will give us a more profound insight into the behaviour of the moiré phenomenon, and lead us to a method for finding moiré-free screen combinations for colour separation.

Let us start with the simple case of two superposed regular dot screens. As we have seen, any of the convolution impulses in the spectrum which fall within the visibility circle represent a moiré phenomenon between the two screens. Let the frequencies of the 2 screens be f_1 and f_2 and the angles they form with the positive horizontal axis be θ_1 and θ_2 . Then according to Eq. (17) the geometric location of the (k_1, k_2, k_3, k_4) -impulse in the spectrum is given by the Cartesian coordinates (f_u, f_v) where:

$$\begin{aligned} f_u &= k_1 f_1 \cos \theta_1 + k_2 f_1 \cos(90^\circ + \theta_1) + k_3 f_2 \cos \theta_2 + k_4 f_2 \cos(90^\circ + \theta_2) \\ f_v &= k_1 f_1 \sin \theta_1 + k_2 f_1 \sin(90^\circ + \theta_1) + k_3 f_2 \sin \theta_2 + k_4 f_2 \sin(90^\circ + \theta_2) \end{aligned} \quad (18)$$

or equivalently, in vector form:

$$\begin{pmatrix} f_u \\ f_v \end{pmatrix}_{k_1, k_2, k_3, k_4} = \begin{pmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} f_1 + \begin{pmatrix} \cos \theta_2 & -\sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 \end{pmatrix} \begin{pmatrix} k_3 \\ k_4 \end{pmatrix} f_2$$

Obviously, if we now let the parameters f_i and θ_i vary, the geometric location of the (k_1, k_2, k_3, k_4) -impulse varies accordingly in the spectrum. For the sake of simplicity we may assume without loss of generality that the first screen is fixed, with the angle $\theta_1 = 0^\circ$ and the frequency of f_1 , and only the second screen is free to vary. Thus, in the superposition of 2 regular dot screens there are only 2 independent parameters (degrees of freedom); it is convenient to choose them to be the angle difference between the 2 screens, $\alpha = \theta_2 - \theta_1 = \theta_2$, and their frequency ratio, $q = f_2/f_1$. Expression (18) for the geometric location of the (k_1, k_2, k_3, k_4) -impulse in the spectrum can be rephrased, therefore, as a function of the 2 variables α and q , as follows:

$$\begin{aligned} f_u &= k_1 f_1 + q f_1 [k_3 \cos \alpha + k_4 \cos(90^\circ + \alpha)] \\ f_v &= k_2 f_1 + q f_1 [k_3 \sin \alpha + k_4 \sin(90^\circ + \alpha)] \end{aligned} \quad (19)$$

The distance of this impulse from the spectrum origin (i.e. the frequency of the impulse) is given, according to Eq. (6), by:

⁵ We will restrict ourselves here only to dynamics of the *geometric* parameters, i.e. to the influence of varying the *angles* and *frequencies* of the superposed layers on the resulting moirés. The effects of varying the gray levels or the dot shapes in the superposed layers have already been discussed above (see Figs. 8, 12, 13).

$$f_{k_1,k_2,k_3,k_4} = \sqrt{f_u^2_{k_1,k_2,k_3,k_4} + f_v^2_{k_1,k_2,k_3,k_4}}$$

and its period is: $T_{k_1,k_2,k_3,k_4} = 1/f_{k_1,k_2,k_3,k_4}$.

Let us take as an example the (1,1,-1,0)-moiré of Fig. 9(b), i.e. the impulse with $k_1=1$, $k_2=1$, $k_3=-1$ and $k_4=0$. As we can see in Fig. 10(b), the geometric location of this impulse (denoted in the figure as the vectorial sum) is quite close to the spectrum origin; and in fact, if we rotate and scale the free screen so that α and q vary towards the values of $\alpha=45^\circ$ and $q=\sqrt{2}$, the impulse will gradually approach the origin, and finally reach it. At this precise point the period of the moiré is infinitely large, but at any values of α and q around this point the impulse in the spectrum will still be located within the visibility circle, so that the moiré will have a large, visible period. It can be said, therefore, that in the 2D *parameter space* defined by all the possible values of the parameters α and q , the point $(\alpha,q)=(45^\circ,1.4142)$ represents a singular state of the (1,1,-1,0)-moiré. Moreover, this is the *only* singular point of the moiré, as a glance at Fig. 10(b) shows: since the first screen is fixed, so is the sum of its two \mathbf{f}_1 vectors; and only at $\alpha=45^\circ$, $q=\sqrt{2}$ the vector $-\mathbf{f}_2$ of the second (varying) screen cancels it out and brings the vectorial sum (the small arrow) to zero. A similar reasoning shows that any (k_1,k_2,k_3,k_4) -moiré between 2 screens has a single singular point in the 2D parameter space.

However, every (k_1,k_2,k_3,k_4) -moiré may have a different singular point in the parameter space. This is demonstrated in Fig. 14, which gives a panoramic view of the parameter space (α,q) and illustrates the dynamics of the most important moirés (up to the second order) which appear between the 2 superposed screens. Note that different moirés may have an identical singular point in the parameter space; this is the case, for instance, in the (1,1,-1,0)- and (2,0,-1,1)-moirés (see Fig. 14). We will call such moirés *isocentric* since they share the same central (singular) point. The contour line around each moiré in the figure delimits the parameter combinations (α,q) for which the fundamental impulse of the moiré is located inside the visibility circle.

Having illustrated the dynamics of moiré effects between two superposed dot screens, we are now ready to examine the case of 3 superposed screens, which is the basic configuration for colour separation.

For the sake of clarity, let us adopt the following notational conventions: we will call the 3 dot screens, in descending order of their frequencies, the K-screen, the M-screen and the C-screen (shorthand for: black, magenta and cyan). We suppose without loss of generality that the K-screen is fixed, with the angle $\theta_1 = 0^\circ$ and the frequency of f_K , and that only the C- and M-screens are free to vary. The angle between the K- and the M-screens will be denoted by α and the angle between the K- and the C-screens by β (see Fig. 15).

The case of 3 superposed screens therefore has 4 independent parameters: the angles α and β , and the frequency ratios $q_{MK}=f_M/f_K$ and $q_{CK}=f_C/f_K$. The geometric location of the $(k_1, k_2, k_3, k_4, k_5, k_6)$ -impulse in the spectrum of the superposition can be expressed (according to Eq. 17) as a function of these 4 parameters, as follows:

$$\begin{aligned} f_{u, k_1, \dots, k_6} &= k_1 f_K + q_{MK} f_K [k_3 \cos \alpha + k_4 \cos(90^\circ + \alpha)] + q_{CK} f_K [k_5 \cos(-\beta) + k_6 \cos(90^\circ - \beta)] \\ f_{v, k_1, \dots, k_6} &= k_2 f_K + q_{MK} f_K [k_3 \sin \alpha + k_4 \sin(90^\circ + \alpha)] + q_{CK} f_K [k_5 \sin(-\beta) + k_6 \sin(90^\circ - \beta)] \end{aligned} \quad (20)$$

$$\text{with } f_{k_1, \dots, k_6} = \sqrt{f_{u, k_1, \dots, k_6}^2 + f_{v, k_1, \dots, k_6}^2} \quad \text{and} \quad T_{k_1, \dots, k_6} = 1/f_{k_1, \dots, k_6}.$$

Since the parameter space in this case is 4-dimensional, $(\alpha, \beta, q_{MK}, q_{CK})$, a full graphic description (like Fig. 14 for the 2D case) is no longer possible. But, except for the abstraction due to the 4 dimensions, the situation remains basically similar. The main difference is that unlike in the 2D case, the locus of the singular points of a $(k_1, k_2, k_3, k_4, k_5, k_6)$ -moiré in the 4D parameter space is no longer a single point, but rather consists of a 2D manifold (a curved surface) within the 4D space. This is illustrated in Figs. 16-17.

As the parameters $(\alpha, \beta, q_{MK}, q_{CK})$ vary away from this singular manifold, the period of the moiré becomes smaller until at a certain distance from the singular manifold the moiré fades out and becomes practically invisible. If we “draw” in the 4D space the locus of all the points $(\alpha, \beta, q_{MK}, q_{CK})$ at which the $(k_1, k_2, k_3, k_4, k_5, k_6)$ -impulse is located inside the visibility circle in the spectrum (i.e. $f_{k_1, \dots, k_6} < \text{visibility circle radius}$), we will get around the singular manifold a 4D zone (which is analogous to the spherical zone around each singular point in the 2D case of Fig. 14). We will call this zone the *forbidden zone* of the moiré in question. In 2D sections taken through the 4D parameter space this forbidden zone may appear as a thick, curved or straight line (Fig. 17(a)), or as an elliptic or egg-like shape (Fig. 17(b)).

Now, if we “draw” in the 4D space the forbidden zones belonging to all the $(k_1, k_2, k_3, k_4, k_5, k_6)$ -impulses in the spectrum which represent perceivable moirés, we will get the following image: the singular manifolds of these moirés will appear dispersed in the 4D space, each of them serving as a skeleton which is surrounded by the forbidden zone of the corresponding moiré. These forbidden zones are sometimes intersected, and sometimes they only almost touch in the 4D space. The spaces left *between* these forbidden zones include the parameter combinations for which no disturbing moiré occurs between the 3 screens.

12. Finding moiré-free screen combinations for colour separation

The above discussion immediately suggests a practical way of searching for moiré-free combinations (or *moiré minimization*) for the C,M,K dot screens in colour separation. In principle, any parameter combination in which no significant moiré is visible is a solution to the problem, including the singular states at the centers of the forbidden zones (unless such a singular point is also covered by a forbidden zone belonging to another moiré). As we have already seen in Sec. 10, the traditional combination of $\alpha=\beta=30^\circ$ and $q_{MK}=q_{CK}=1$ is an example of such a singular point. But when the accuracy of the printing process is not enough to guarantee the high precision required for the use of a singular screen combination, it may be preferable to use a stable moiré-free solution, i.e. a screen combination which is located in the 4D space at the center of a moiré-free zone, with as high as possible tolerances in each of the 4 parameters involved. Therefore, rather than looking for singular points (in the center of the forbidden zones), as proposed in [14], we will be looking for stable solutions in the center of the moiré-free zones.

Before we proceed to describe the algorithm itself, a word about the useful ranges of the 4 parameters is in order. It is clear from Fig. 15 that the angles α and β are limited to the range $0^\circ \leq \alpha+\beta \leq 90^\circ$, but since we know experimentally that any angle difference lower than 20° inevitably generates strong {1,1}-moirés (such as in Fig. 9(a)) we may limit ourselves only to cases in which all of the 3 angles, α , β and γ , are larger than 20° . This implies the following angle range (see Fig. 18): $20^\circ \leq \alpha \leq 50^\circ$, $20^\circ \leq \beta \leq 70^\circ - \alpha$. As for the frequency ratios, since we suppose that $f_C f_M \leq f_K$ it follows that the useful ranges for q_{MK} and q_{CK} are: $q_{CK}, q_{MK} \leq 1$. But since it is inappropriate to mix within the same color image C,M,K screens of very different periods or frequencies we will restrict ourselves to the following period ranges: $T_K \leq T_M, T_C \leq 1.5 * T_K$, which means in terms of frequency ratios: $2/3 \leq q_{CK}, q_{MK} \leq 1$.

The method we adopted for finding moiré-free combinations of screens for colour separation is a straightforward search, which runs with the specified steps through all the parameter combinations, within their admitted ranges. It first searches for good combinations for the C,M,K screens, and only then, for each good solution it attempts to find appropriate parameters for the Y screen. Here are the main steps of the algorithm:

Step 1: Generating a list of the “dangerous” impulses in the spectrum of the superposition.

The aim of this step is to select from the list of impulses generated in the spectral convolution of the C,M,K screens a relatively small subset which contains only the

“potentially dangerous” impulses. This will be further explained below. It should be noted that what is needed here is not the actual impulse locations or amplitudes, but just their indices, such as: (2,0,-1,0,0,1).

Step 2: Scanning the 4D parameter space for moiré-free combinations.

In this stage we run through all combinations of the 4 parameters α , β , q_{MK} , q_{CK} (within their admitted ranges, as described above). The step size should be small enough to ensure that no moiré-free zones escape through the mesh, but not too small either in order to avoid excessive running time. For each parameter combination the program runs through the list of “dangerous” impulses and calculates for each of them its actual coordinates in the spectrum, using Eq. (20). If any of the impulses fall inside the corresponding visibility circle, the current parameter combination is rejected. Only if all the dangerous impulses fall outside their visibility circle is the parameter combination accepted and registered in a file which accumulates the potentially good moiré-free combinations.

Step 3: Finding the best solutions inside the detected moiré-free zones.

Having completed the scanning of all the 4 parameter combinations in step 2, we are left with a file containing all the potentially good combinations found. As we have seen above, these points are located within the moiré-free zones in the 4D parameter space. In this step we search for the best solutions among these points, according to the tolerances they offer. Then each of the best solutions found must be experimented on high resolution films, in order to exclude the possible existence of residual moirés which are still visible but have not been detected by the program (for instance, higher order moirés).

Step 4: Finding for each good C,M,K combination an appropriate Y screen.

After the proposed moiré-free combinations for the C,M,K screens have been tested and verified as good solutions, a similar approach is taken to find for each of them a good Y screen. In this case the values of α , β , q_{MK} and q_{CK} are already known, and the varying parameters are the angle δ and the frequency ratio q_{YK} of the Y screen. (Obviously, the list of “dangerous impulses” as well as Eq. (20) must be adapted to the case of 4 superposed screens).

The reason why the best Y screen is found separately is double: first, the introduction of a fourth screen in steps 1–3 would significantly increase the number of moirés which occur between the different layers, and make it practically impossible to find good moiré-free combinations between all 4 layers simultaneously. As a compromise, it is better to find the best possible moiré-free combinations for the C,M,K screens first, even at the expense of having somewhat stronger moirés with the Y screen (which are

anyway much less visible due to the light nature of the Y colour). As a second benefit, the separation of the Y screen from the others reduces from 6 to 4 the number of dimensions to be scanned in step 2; this significantly reduces the complexity of the program, particularly in terms of running time.

Let us return now in more detail to step 1, concerning the number of impulses involved. Since each of the C,M,K screens contributes 2 perpendicular Dirac-combs to the convolution, the total number of impulses in the convolution, if only N impulses (harmonics) are considered from each side of the DC, is $(2N+1)^6$. For $N=2$ this gives $5^6=15625$ impulses, and for $N=3$ the impulse number becomes as high as $7^6=117649$. Fortunately, however, the actual number of impulses to be considered can be significantly reduced by means of 3 different considerations:

First, according to our practical experience, an impulse generated by a harmonic higher than 2 very rarely contributes a significant visible moiré. It is therefore more efficient to search for good solutions taking into account only 2 harmonics, and only then test each of the received solutions for possible higher order moirés.

Second, since we only deal with real images having no imaginary part, each impulse in the spectrum is accompanied by a twin impulse, symmetric with respect to the DC (see Sec. 2). Judicious consideration enables us to leave in the impulse list only one impulse from each such pair, thus reducing by half the number of impulses in the list.

Third, even after the above cut-downs in the number of impulses, most of the impulses remaining in the list have no practical significance from the moiré point of view since they represent moirés which are far too weak to be perceived, even when they appear at low frequencies (for instance, such are all the impulses having more than 2 indices with the value of ± 2). After discarding all these negligible impulses, we are left with a list of only about 2000 “potentially dangerous” impulses, i.e. impulses which may represent, when they fall inside the visibility circle, significant moirés that should be rejected.

However, as we already know, even these 2000 impulses are not all of the same importance. For example (see Fig. 9), a (1,0,-1,0,0,0)-moiré is much stronger than a (1,2,-2,-1,0,0)-moiré, and consequently its cutoff frequency (i.e. the frequency at which it becomes invisible to the eye) is higher. In fact, strong moirés may still cause a rough and clearly visible linen-like structure even at very small periods, where the weaker moirés already disappear completely. Therefore the 2000 impulses in the list have been classified into several categories, according to the relative strength of their moirés. Each category has been assigned its own visibility circle, according to its cutoff frequency. The cutoff frequency for each type of moiré has been found experimentally by measurements performed on a light-table, using various high-resolution films with different screen frequencies, cell shapes and gray levels. In each case the observer

looked at the corresponding screen combination from a normal reading distance, and he rotated the films slightly so as to decrease the period of the moiré in question (i.e. to increase its frequency) until, at a certain moment, the moiré completely disappeared. The highest visible frequency of the moiré, just before its disappearance, was measured and taken as its cutoff frequency (allowing for some extra safety margins). Moiré types with similar cutoff frequencies have been grouped into the same category. In our specific implementation we had 4 impulse categories: one for the strongest moirés (of the $\{1,1\}$ type), a second one for the slightly weaker moirés of the $\{1,1,1\}$ type, a third category for the medium-strength moirés such as $\{1,1,2\}$ and $\{1,2,2\}$, and the last category for all the weaker cases still in the list. It is interesting to mention that if no classification of the impulses is carried out and all of them use the visibility circle of the strongest moirés, the forbidden zones they generate in the 4D parameter space become larger and cover the whole of the space, so that no solutions can be found.

Note that these experimental measurements of the cutoff frequency have the advantage of being performed using real halftone screen superpositions, thus taking into account the irregular background and the masking effects which influence the moiré visibility. This is much better than using standard data taken from literature, which have been measured on ideal sinusoidal images, free of noise such as rosettes or textures due to other moirés in the background.

13. Results and discussion

An important advantage of the above moiré minimization method is that, due to the relative simplicity of the calculations involved, it is capable of scanning the full admitted ranges of the 4 parameters. This provides a global, panoramic view of the whole 4D parameter space, including the forbidden zones inside it and the moiré-free zones between them. Note that every large enough moiré-free zone should be detected by this scan; moiré-free zones which “escape” through the scanning mesh (the sampling points in the 4D space) are inevitably smaller than the scanning step, which means that their tolerance margins are very low so that anyway they do not correspond to our requirements. (Such are, for example, all the singular points). Obviously, after an important moiré-free zone is discovered, the program may be used again to rescan the interesting zone alone in greater detail (“zoom in”).

The results obtained by this method are quite interesting. Several good screen combinations have been found, which gave in experimental tests a very satisfactory, uniform appearance with no evident apparition of moirés or other disturbing structures (some of these solutions correspond to previously known cases listed in [15]). An example of such a good solution is shown in Fig. 19. Note that each solution in fact represents a whole family of equivalent screen combinations which can be received from each other by various symmetries and transformations (such as mirror-images,

rotations, scalings, flipping the colour labels between the 3 screens, etc...). The tolerance obtained around a typical good solution is in the order of at least $\pm 0.5^\circ$ for the angles α and β and $\pm 1\%$ for the screen periods or frequencies (these values are indicated for the narrowest dimension of the moiré-free zone; in its more prolonged directions the tolerances may be even higher). This is significantly better than the tolerance required by the traditionally used singular screen combination of $\alpha=\beta=30^\circ$ and $q_{MK}=q_{CK}=1$ (or any other singular case), which is $\pm 3'$ (minutes of arc) for the angles (DIN No. 16547). And moreover, since our good solutions are not singular states, even when the specified tolerances are somewhat exceeded, the visual effect only consists of weak, small-period moirés which gradually start to appear, while in singular moiré-free cases the slightest deviations immediately bring back from infinity large-period, highly objectionable moirés. In other words: our solutions are *stable* moiré-free states rather than *unstable* moiré-free states.

Another interesting result observed is that the best solutions in the moiré-free zones, according to our criteria of maximal tolerance, were systematically located on a symmetry axis of the 4D parameter space, such as: $\alpha=\beta$, $q_{MK}=q_{CK}$ (or any of its symmetric counterparts, like: $\beta=\gamma$, $q_{CK}=q_{CM}$ etc.). Perhaps this is not so surprising, after all, since there is no a-priori reason why one side of the symmetry axis should be preferred to the other. But the interesting thing is that these highly symmetric cases were not necessarily the winners during the experimental tests, and often the cases found most suitable for practical use were slightly *off* the symmetry axes. The reason for this result appears to be the following: in highly symmetrical cases the dots of the 3 screens form together in the superposition a fine uniform geometrically ordered micro-structure which appears very regular to the eye. Now, even very weak moirés (including moirés due to higher harmonic orders!), which are normally completely negligible on the irregular, noisy background of the screen superposition (even if their period is relatively large), may become quite disturbing here. Even the slightest perturbation on a uniform, geometrically ordered background is very clearly perceived by the human eye which is very sensitive to such periodic irregularities on a uniform background.

Therefore, even if the most promising cases obtained by the program are located on a symmetry axis, it might be advisable to prefer a neighbouring case slightly off that axis, in order to break the symmetry of the micro-structure generated in the superposition.

14. Conclusions

In this article a spectral approach for the analysis of superposition moiré effects is presented. Using this approach, the development of the moiré analysis is systematically surveyed starting from the simplest, cosinusoidal grating case, through binary gratings

and square grids to the case of dot screen superposition, and then further to the case of superposition moirés in colour separation. This is done based on a useful notational formalism introduced for the identification, classification and labelling of the moiré effects. Interesting results which have been demonstrated include the fact that the angles and periods of the moiré patterns are determined only by the angles and periods of the superposed structures, while the intensity and the profile shape of the moiré patterns are basically determined by the opening ratios or the dot shape of the superposed structures. We also define the visibility circle for each type of moiré, the radius of which depends on several factors, including the perceptual contrast of the moiré.

Then, based on this background, in the second part of the article we analyze the moiré phenomenon from a different angle — the dynamic point of view. We introduce the moiré parameter space and show that the good parameter combinations are located in it within isolated moiré-free zones. This leads us to an algorithm for finding stable moiré-free screen combinations for colour separation, which have much better tolerances than singular moiré-free screen combinations.

Acknowledgement

This research was partially supported by the Swiss National Fund (grant No. 21-31136.91).

References:

1. G. Oster, *The Science of Moiré Patterns*. Edmund Scientific Co., USA (second edition, 1969).
2. P. S. Theocaris, *Moiré Fringes in Strain Analysis*. Pergamon Press, UK (1969).
3. H. Takasaki, "Moiré topography," *Applied Optics*, **9**(6), 1467–1472 (1970).
4. Y. Nishijima and G. Oster, "Moiré patterns: their application to refractive index and refractive index gradient measurements," *Jour. of the Optical Society of America*, **54**(1), 1–5 (1964).
5. J. A. C. Yule, "Moiré Patterns and Screen Angles," Chap. 13 in *Principles of Color Reproduction*, pp. 328-345, John Wiley & Sons, Inc., New York (1967).
6. I. Amidror, "The moiré phenomenon in colour separation," in *Raster Imaging and Digital Typography II*, Proc. of the 2nd International Conference on Raster Imaging and Digital Typography, R. A. Morris, J. André, Eds., pp. 98–119, Cambridge University Press (1991).
7. G. Oster, M. Wasserman and C. Zwerling, "Theoretical interpretation of moiré patterns," *Jour. of the Optical Society of America*, **54**(2), 169–175 (1964).
8. A. Rosenfeld and A. C. Kak, *Digital Picture Processing, Vol. I*. Academic Press, Inc., Florida, USA (second edition, 1982).
9. R. N. Bracewell, *The Fourier Transform and its Applications*. McGraw-Hill Publishing Company, Reading, N.Y. (second edition, 1986).
10. R. W. G. Hunt, *The Reproduction of Colour in Photography, Printing and Television*. Fountain Press, England (fourth edition, 1987).
11. O. Bryngdahl, "Moiré and higher grating harmonics," *Jour. of the Optical Society of America*, **65**(6), 685–694 (1975).
12. W. K. Pratt, *Digital Image Processing*, John Wiley & Sons, Inc., New York (second edition, 1991).
13. W. F. Schreiber, *Fundamentals of Electronic Imaging Systems*. Springer-Verlag, Berlin (1986).
14. P. Fink, *PostScript Screening: Adobe Accurate Screens*, Adobe Press, USA (1992); Chapters 9-11 and Appendix A, pp. 107-146.
15. J. Schoppmeyer, "Screen systems for multicolor printing", U.S. Patent 4537470 (1985).

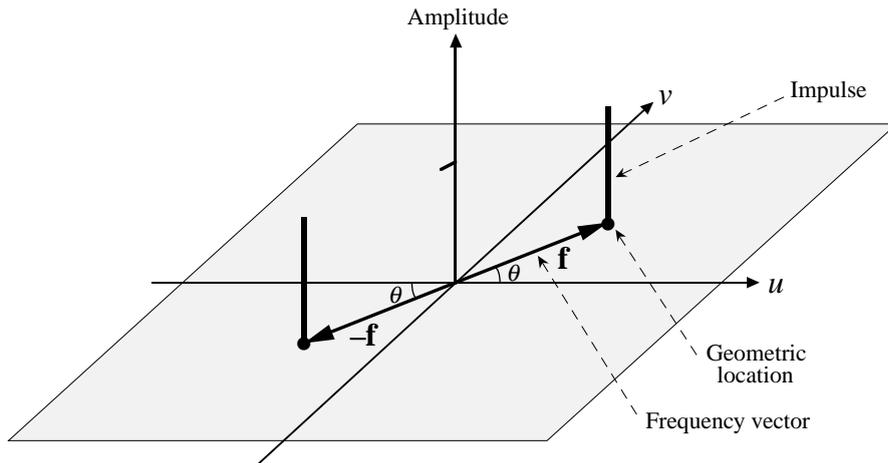


Fig. 1: The geometric location and amplitude of impulses in the 2D spectrum. To each impulse is attached its frequency vector, which points at the geometric location of the impulse in the spectrum plane (u,v) .

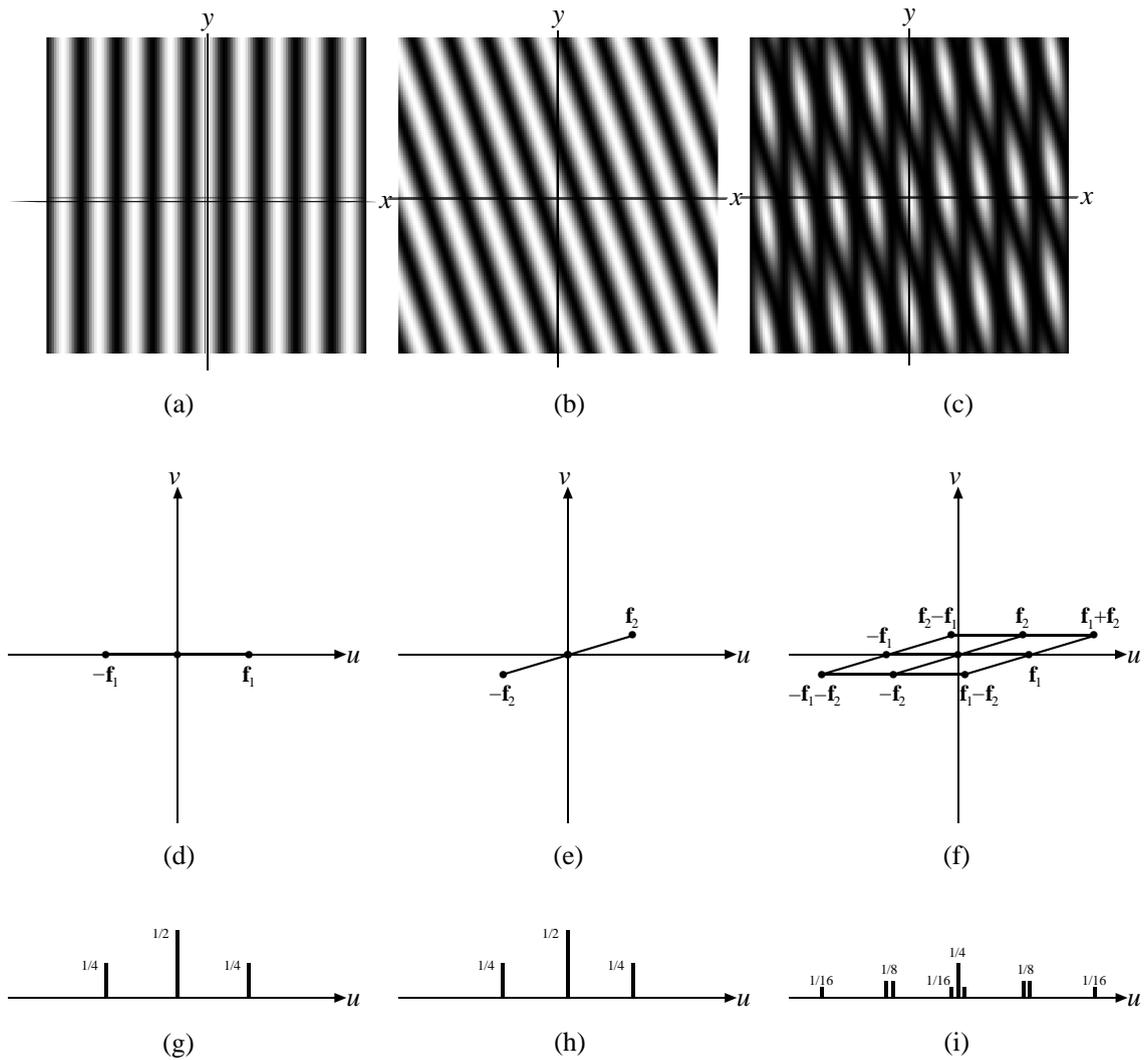


Fig. 2: (a) and (b) Cosinusoidal gratings and (c) their superposition in the image domain. (d) and (e) Top views of the respective spectra and (f) their convolution. Black dots in the spectra indicate the geometric location of the impulses; the line segments connecting them have been added only to clarify the geometric relations. (g), (h) and (i) Side views of the same spectra, showing the impulse amplitudes.

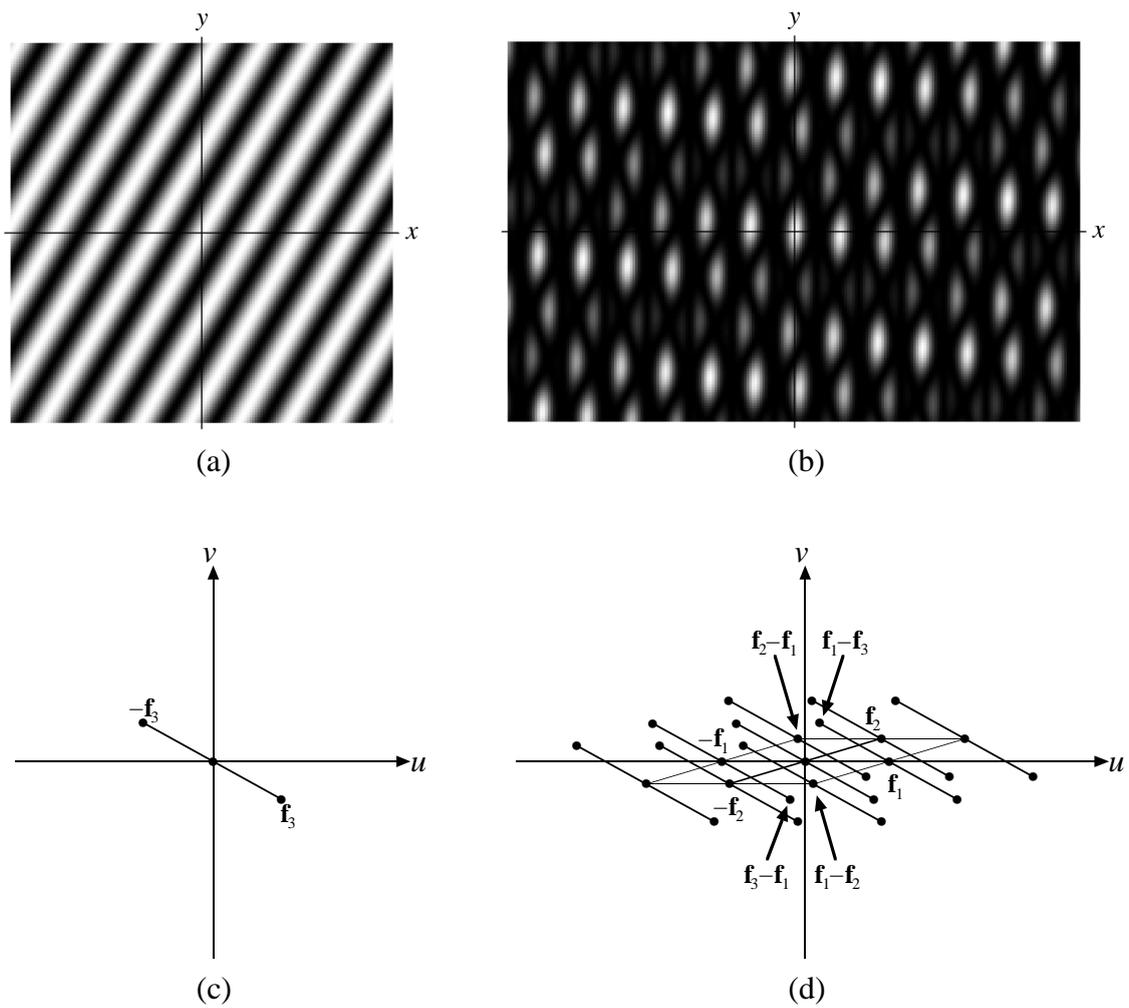


Fig. 3: (a) A third sinusoidal grating and (b) its superposition with Fig. 2(c), in the image domain; (c) and (d) a top view of their respective spectra in the spectral domain. The strongest visible moirés in (b), which are best seen from a distance of about 3 m, belong to the two impulse pairs marked by arrows in the spectrum (d).

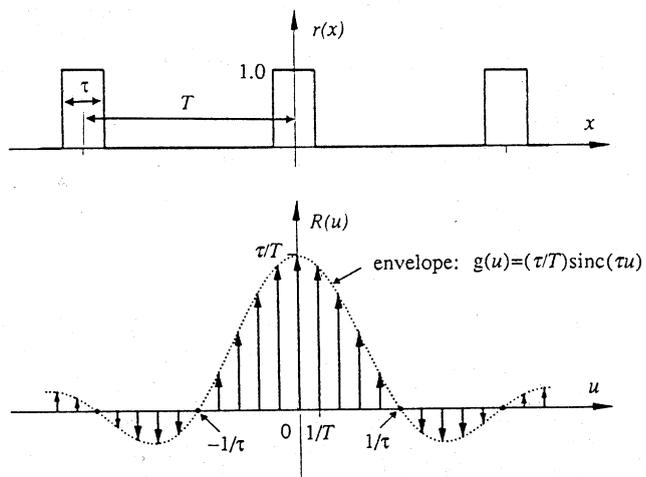


Fig. 4: A symmetric square wave with period T and opening τ and its Fourier transform. The dotted line indicates the envelope of the impulse train.

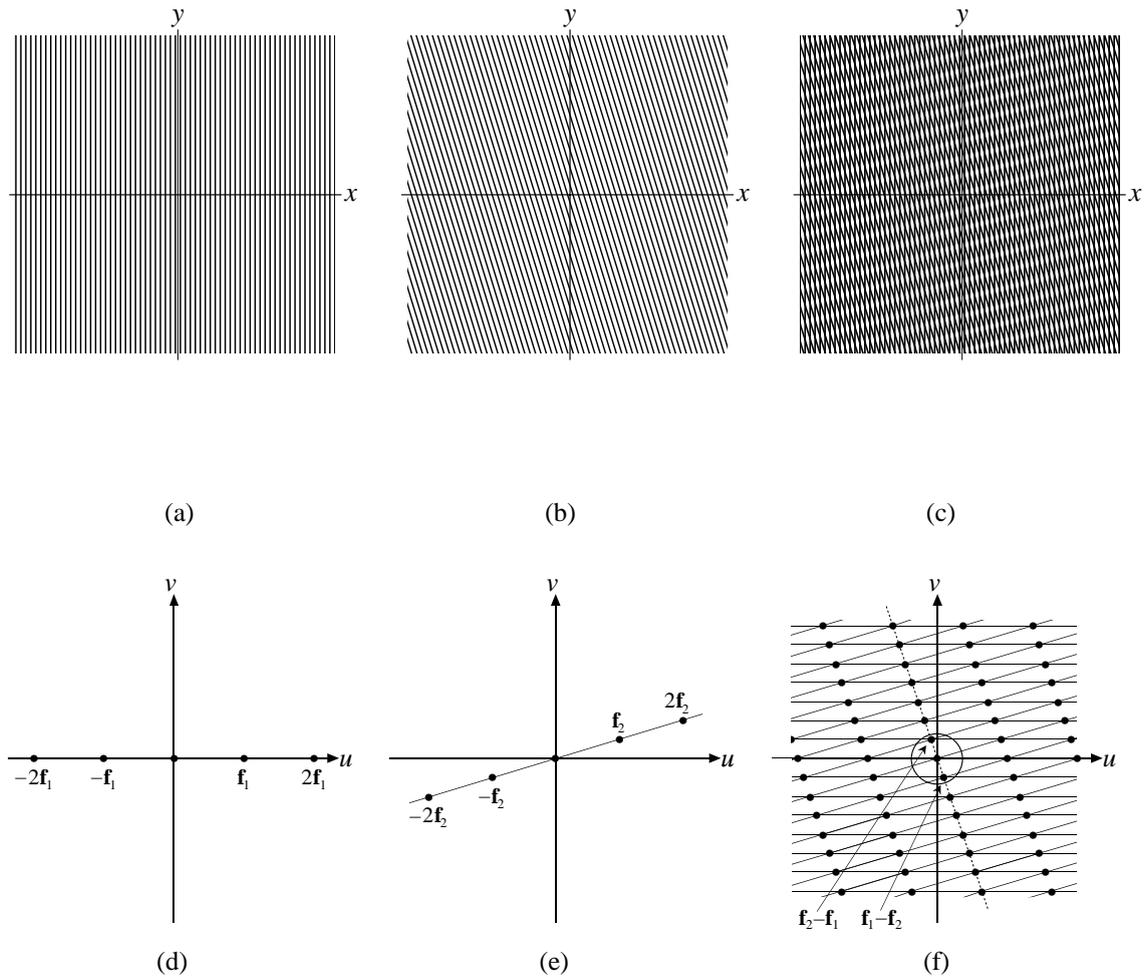


Fig. 5: (a) and (b) Binary gratings and (c) their superposition in the image domain; their respective spectra are the infinite combs shown in (d) and (e) and their convolution (f). (The scale in the image domain was reduced for the sake of clarity). The circle in the center of spectrum (f) represents the visibility circle. It contains the impulse pair whose frequency vectors are $\mathbf{f}_1 - \mathbf{f}_2$ and $\mathbf{f}_2 - \mathbf{f}_1$; this is the fundamental impulse pair of the moiré seen in (c). The dotted line in (f) shows the infinite comb of impulses that represents the moiré.

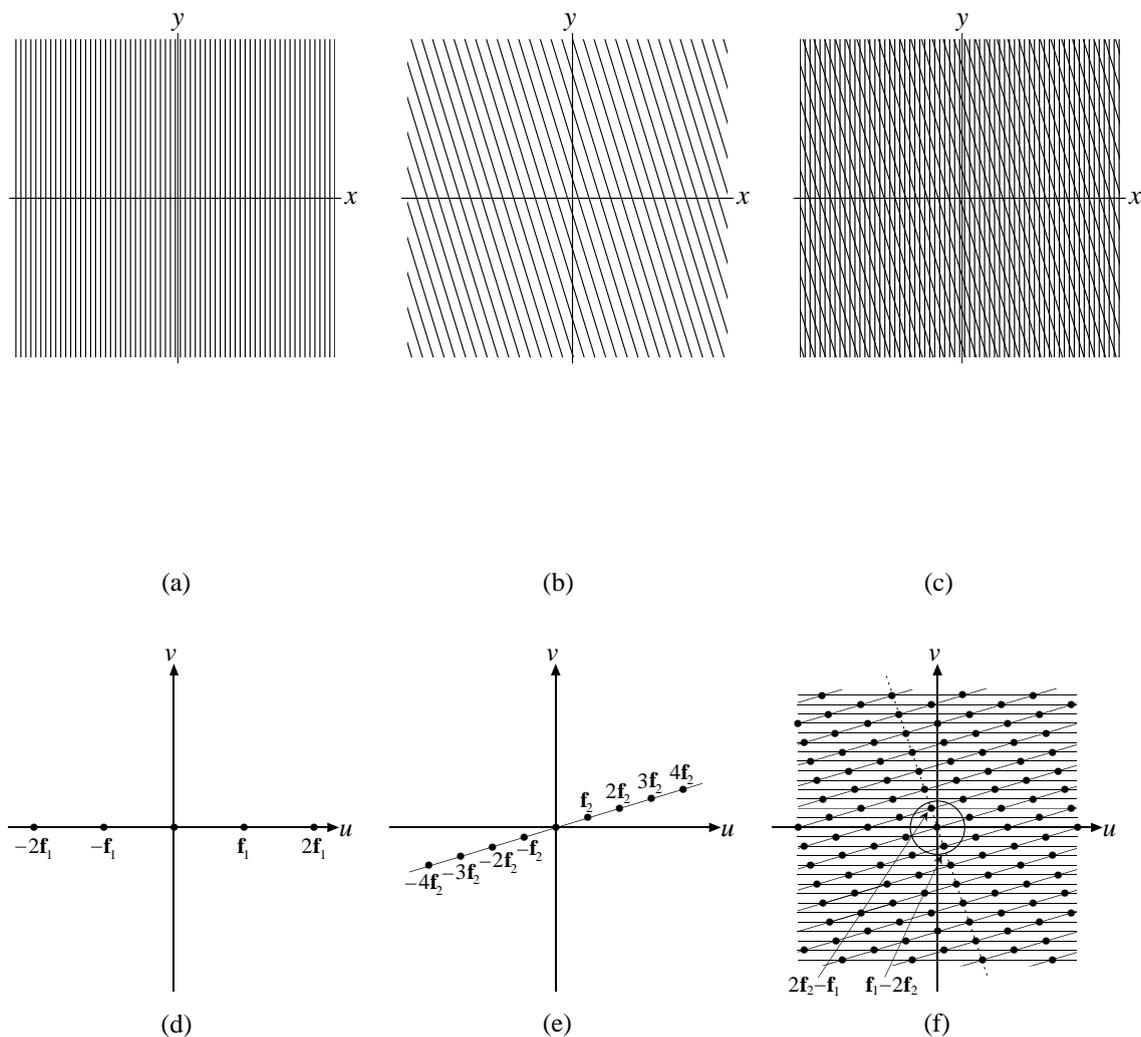


Fig. 6: (a) and (b) Binary gratings as in Fig. 5 but with (b) having half the frequency, and (c) their superposition; (d), (e) and (f) their respective spectra. (The scale in the image domain was reduced for the sake of clarity). The visibility circle in the center of the spectrum (f) contains the impulse pair with frequency vectors $\mathbf{f}_1 - 2\mathbf{f}_2$ and $2\mathbf{f}_2 - \mathbf{f}_1$, which originate from the second harmonic of \mathbf{f}_2 , and represent the fundamental impulse pair of the moiré. Note that the moiré seen in (c) is a (1,-2)-moiré, but it still has the same angle and frequency as the (1,-1)-moiré of Fig. 5, and only its intensity is weaker.

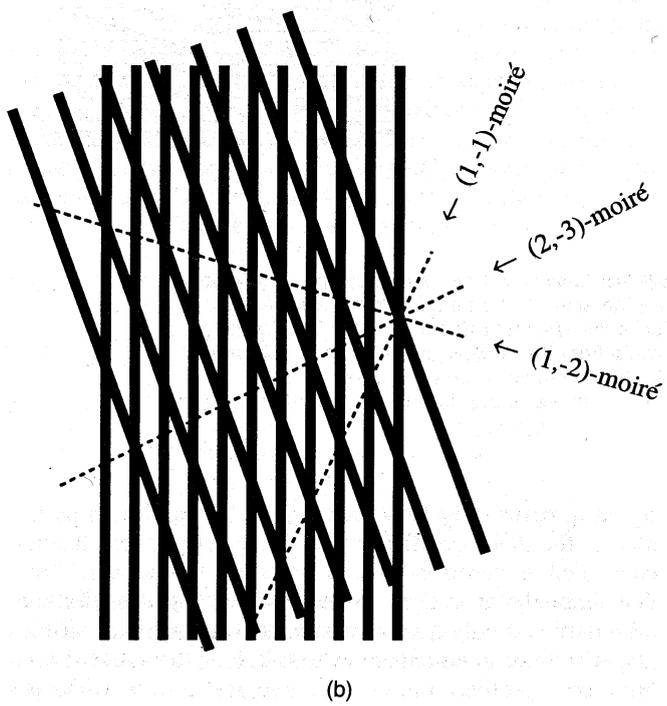
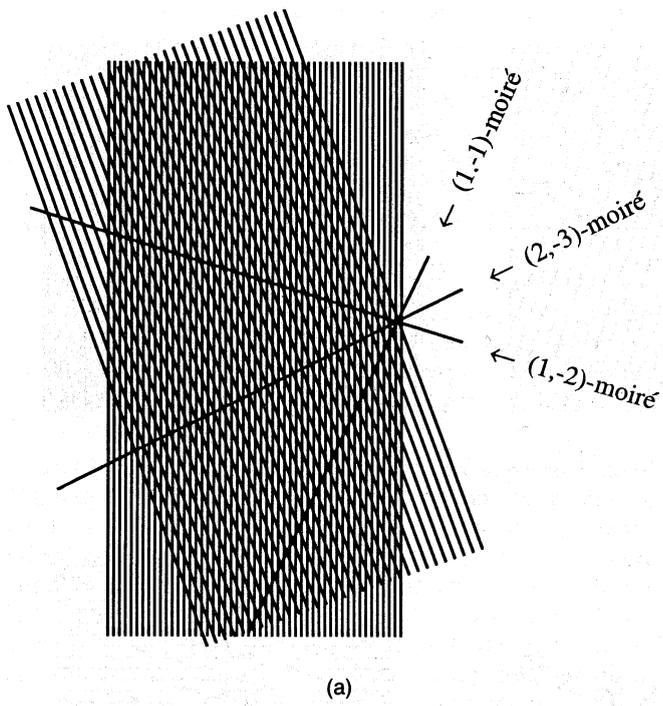


Fig. 7: (a) Some of the different order moirés between two superposed gratings, and (b) enlarged view.

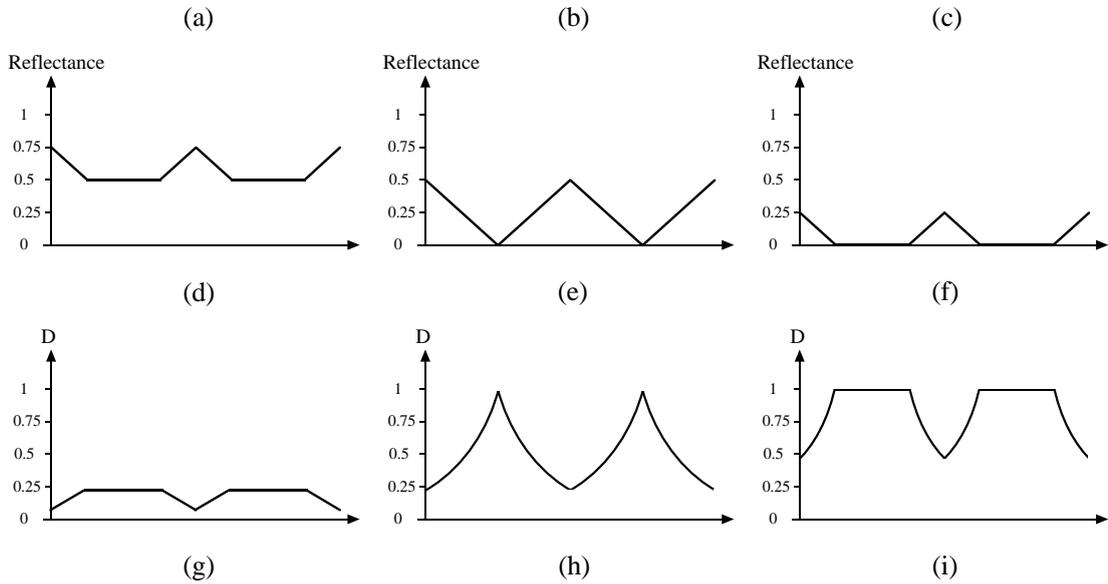
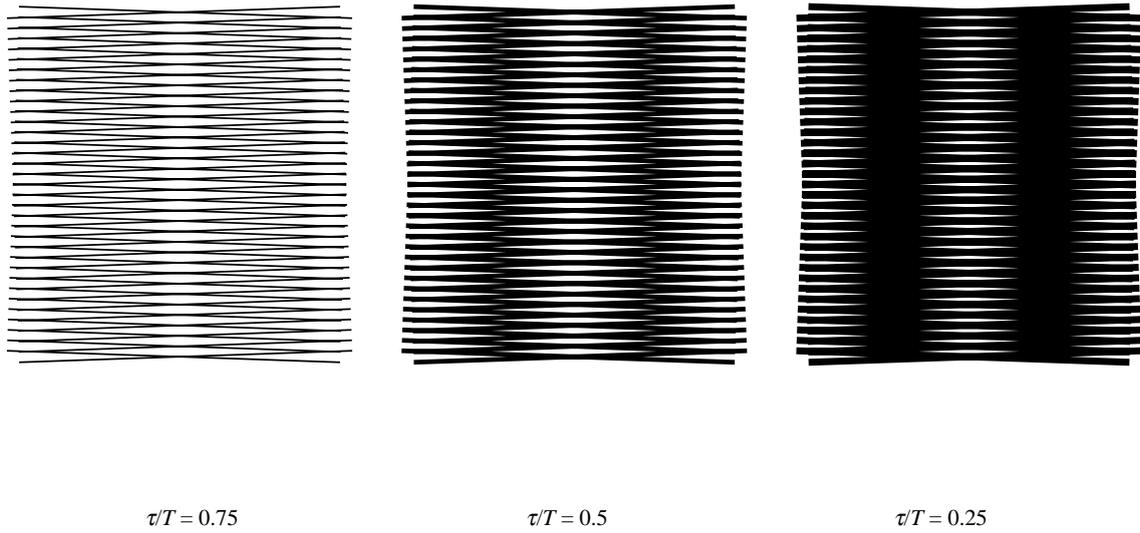


Fig. 8: A (1,-1)-moiré between two identical binary gratings. All three cases have the same periods and angles, but their opening ratios are (a) 0.75, (b) 0.5, and (c) 0.25. (d), (e) and (f) The respective moiré profiles in terms of *reflectance*, as received from the mathematical model. (g), (h) and (i) The same moiré profiles after their adaptation to the human visual perception, i.e., in terms of *density*.

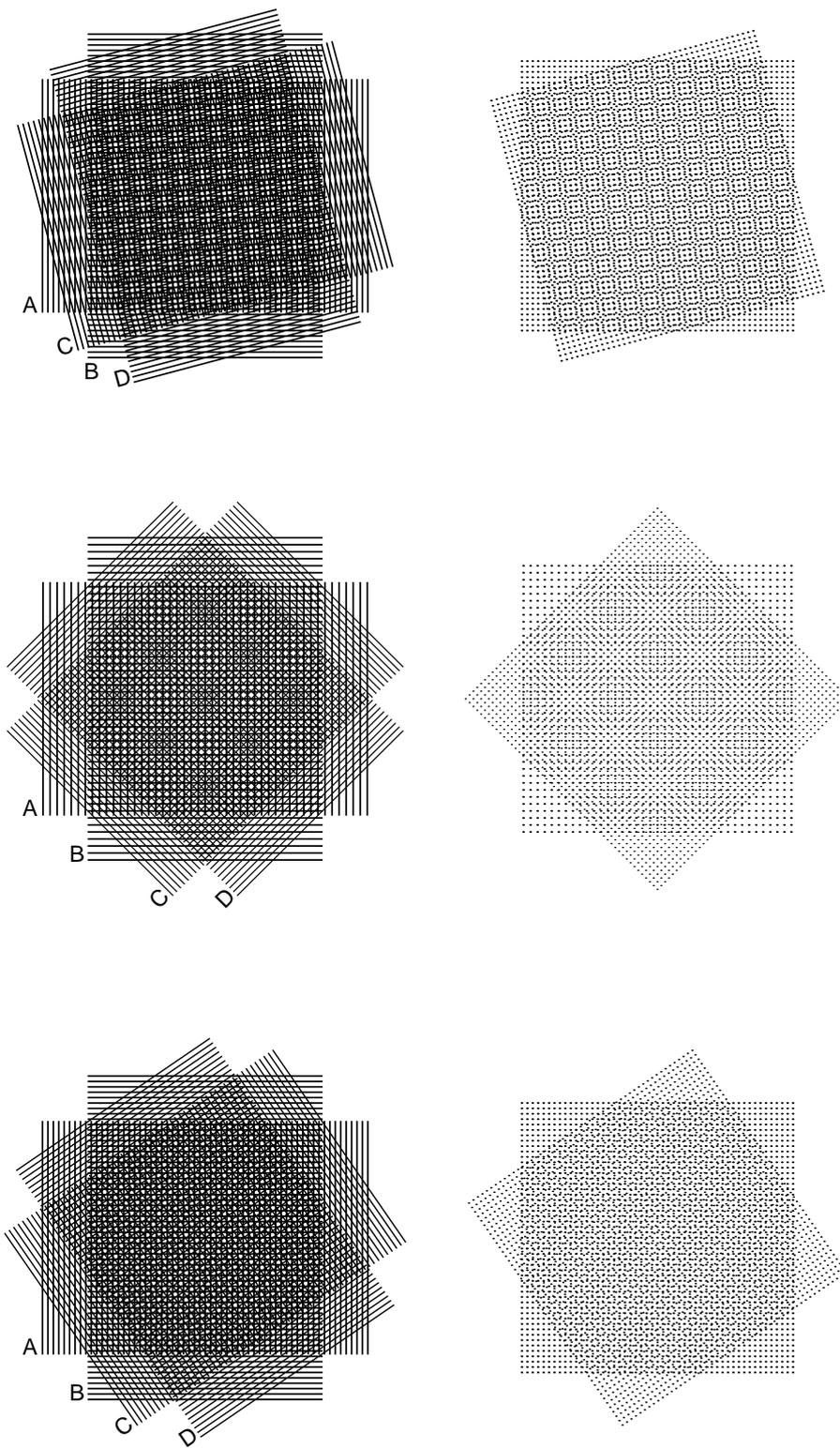


Fig. 9: Three types of moirés between two square grids (left) or two dot-screens (right), which are generated by (a) two, (b) three, or (c) four of the gratings involved.

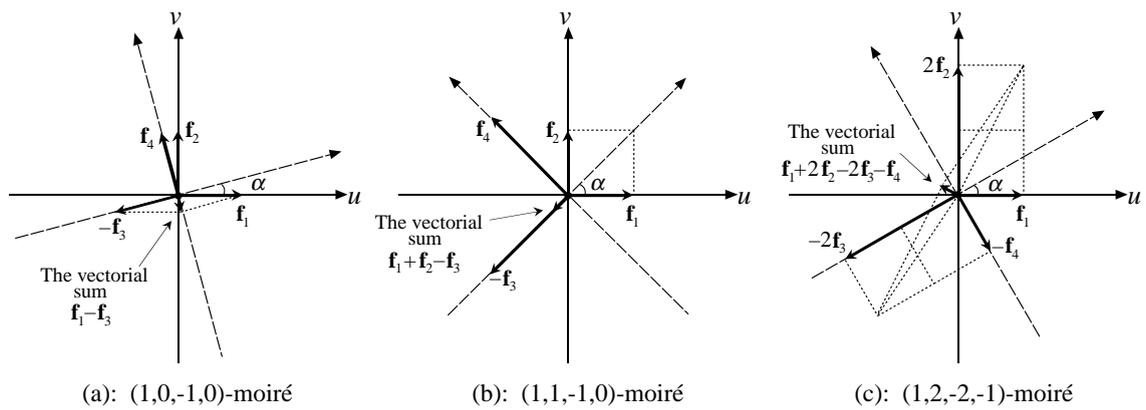


Fig. 10: The spectral interpretation of the three types of moirés between two square grids shown in Fig. 9, which are generated by two, three or four of the gratings involved. For the sake of clarity, only the frequency vectors of one of the two perpendicularly symmetric moirés are shown in each case. The low frequency vectorial sum which corresponds to the fundamental impulse of the moiré in each of these cases can be graphically found by the parallelogram law (the dashed axes here denote the rotated layers).

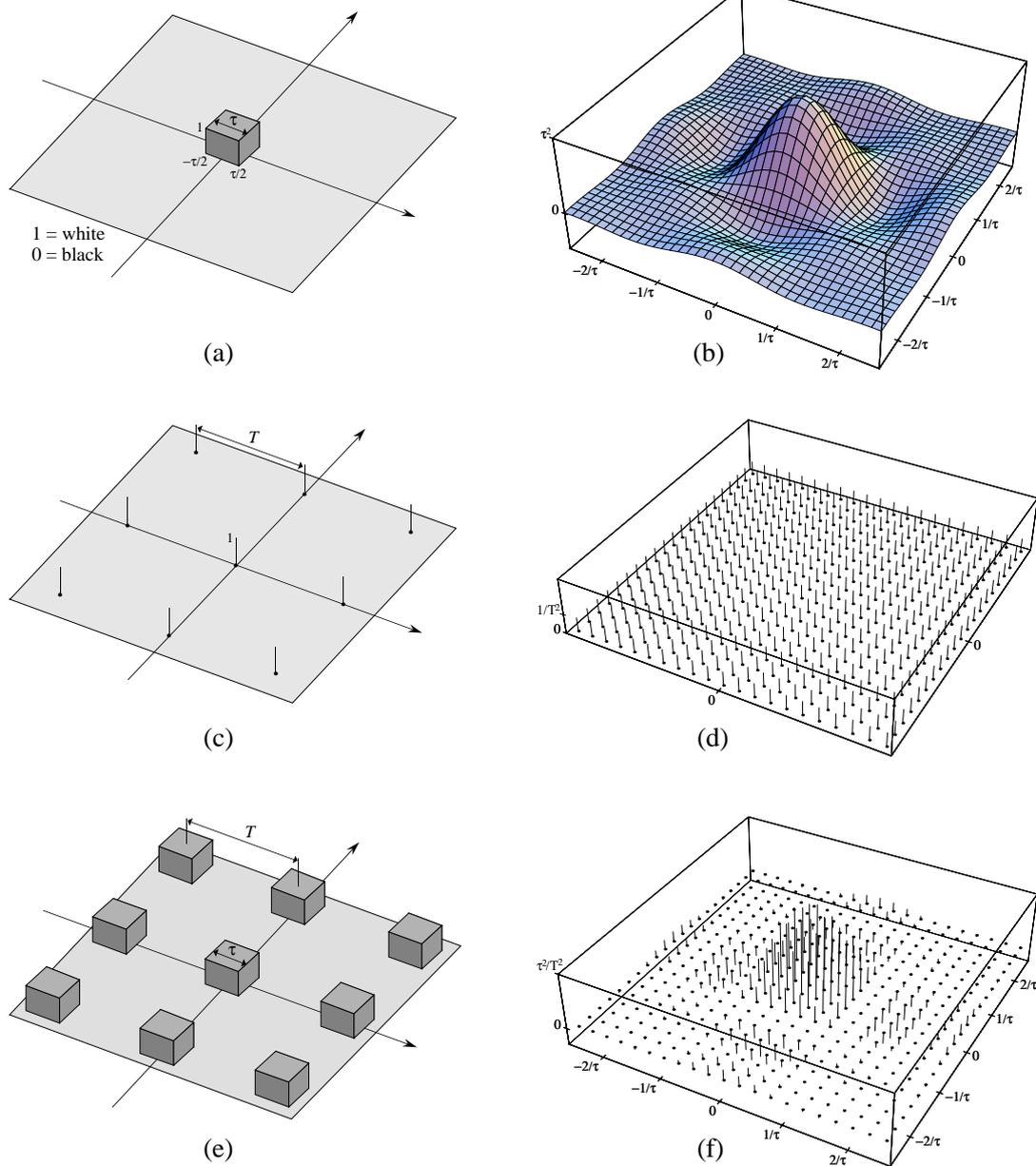


Fig. 11: (a) A square white dot $d(x,y)$ with side τ ; (b) its continuous spectrum $D(u,v) = \tau^2 \text{sinc}(\tau u) \text{sinc}(\tau v)$; (c) a nailbed with period T and amplitude 1; (d) its spectrum — a nailbed with period $1/T$ and amplitude $1/T^2$; (e) a screen of square white dots — the convolution of (a) and (c); and (f) the spectrum of the screen — the product of (b) and (d): a nailbed that samples the “envelope” (b) at intervals of $1/T$, scaling its amplitude by $1/T^2$.

Fig. 12: The influence of the size and shape of the screen dots on the moiré between two superposed screens. Only the size and the shape of the “spot function” of the moiré are affected, but not the period and the direction of the moiré. In both images two similar screens, with identical frequencies and gradually increasing dots, are superposed at an angle difference of 2° . (a) The superposition of two screens with increasingly large circular dots gives circular moiré spots of increasing size. (b) The superposition of two screens with increasingly large square dots give square moiré spots of increasing size.

Fig. 13: The influence of the dot size and of the dot shape of the superposed screens on the moiré intensity. Both images show the superposition of two screens, one at 28.3° and with frequency f_1 and the other at 51.34° and with frequency $f_2 = 1.153 f_1$. In both images the gray level of the first screen is varied vertically from 0 to 1 (by modifying the dot size within the fixed screen period), and the gray level of the second screen varies horizontally from 0 to 1. This means that each image contains all the possible dot size combinations of the two screens. The only difference between images (a) and (b) is in the shape of the screen dots: the dots are circular in (a), and rhombic (i.e., squares rotated by 45°) in (b). In each of the images two distinct moirés appear: The large square pattern at -10.4° , which belongs to a $\{1,2,2\}$ -moiré, and the small square pattern at 20.6° , which is a $\{1,1\}$ -moiré. Note that the intensity of each of the two moirés within each of the images varies with the gray-level combination of the two screens, and reaches its maxima in different gray-level combinations. The influence of the dot shape is demonstrated by the significant difference in the moiré intensities between the two images.

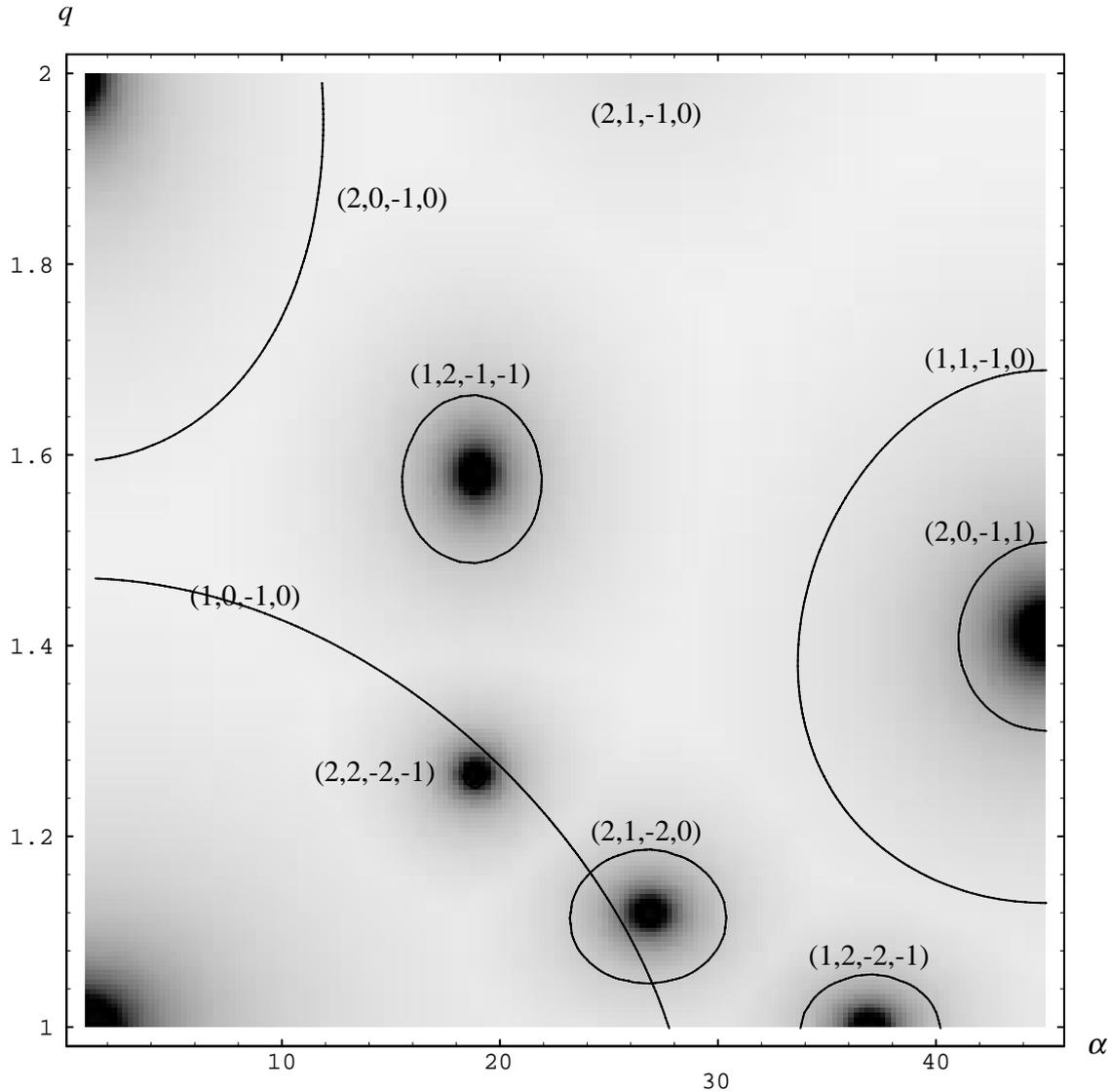


Fig. 14: A panoramic view of the 2D parameter space (α, q) within the range $0^\circ \leq \alpha \leq 45^\circ$, $1 \leq q \leq 2$. It shows the most significant moirés between two superposed regular screens in this range, and illustrates how their period varies as a function of the angle difference α and the frequency ratio q between the two screens. Darker shades represent bigger moiré periods. Each moiré is centered at a *singular point* in which its period is infinitely large; this point is surrounded by a spherical zone in which the period of the moiré gradually decreases from the center outwards. The contour line around each moiré delimits the parameter combinations (α, q) for which the fundamental impulse of the moiré is located in the spectrum inside the visibility circle. Note that the $(1,0,-1,0)$ -moiré is clearly more dominant than the others. Note also that the $(2,2,-2,-1)$ -moiré is already too weak to be visible; in practice it is assimilated by the residues of the strong $(1,0,-1,0)$ -moiré.

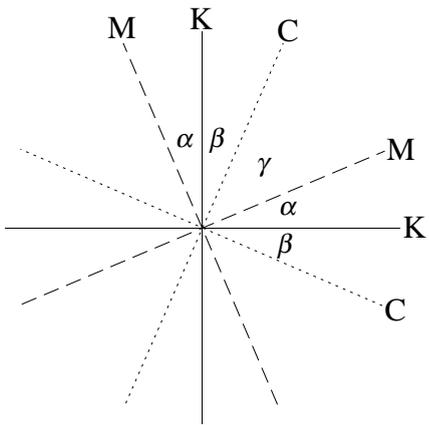


Fig. 15: A schematic view of the three superposed C,M,K screens (each represented by its two main axes) showing the conventional notations used in the text.

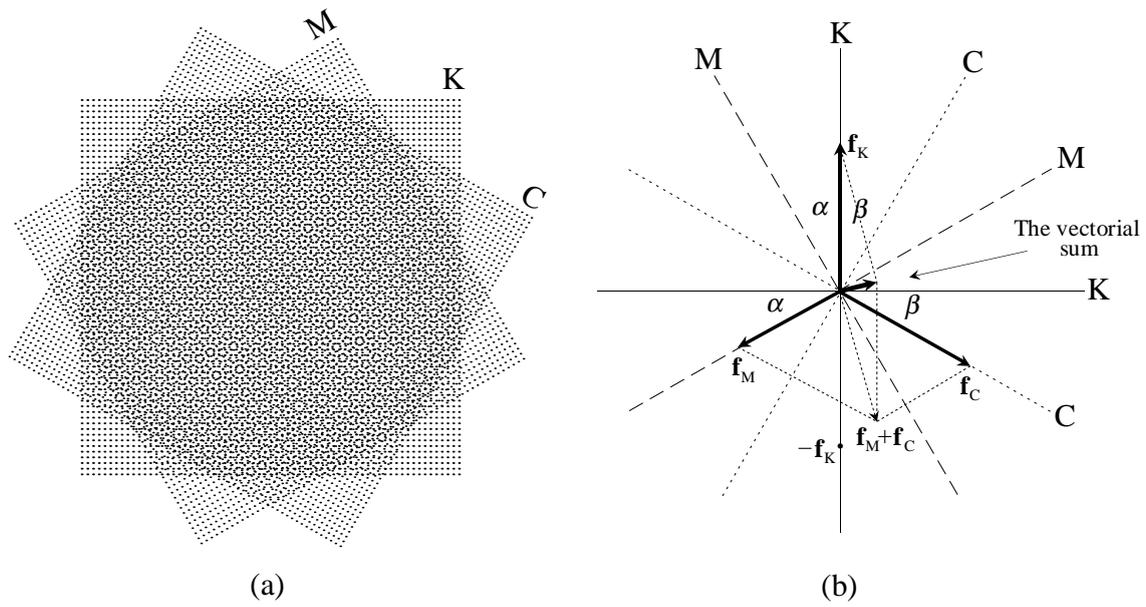


Fig. 16: A $(0,1,-1,0,1,0)$ -moiré between three screens (a), and its vector diagram (b). By convention it is assumed that the K screen is fixed, and only the M and C screens are free to vary. For any choice of α and β there exists one set of q_{MK} and q_{CK} (i.e., one set of \mathbf{f}_M and \mathbf{f}_C) for which $\mathbf{f}_M + \mathbf{f}_C$ falls on $-\mathbf{f}_K$, thus yielding a singular case (where the vectorial sum is $\mathbf{0}$). This means that the locus of all the singular points of this moiré forms a 2D manifold in the 4D space $(\alpha, \beta, q_{MK}, q_{CK})$. Note that one of these singular points, the point with $\alpha = \beta = 30^\circ$ and $q_{MK} = q_{CK} = 1$, is the case used in traditional colour printing.

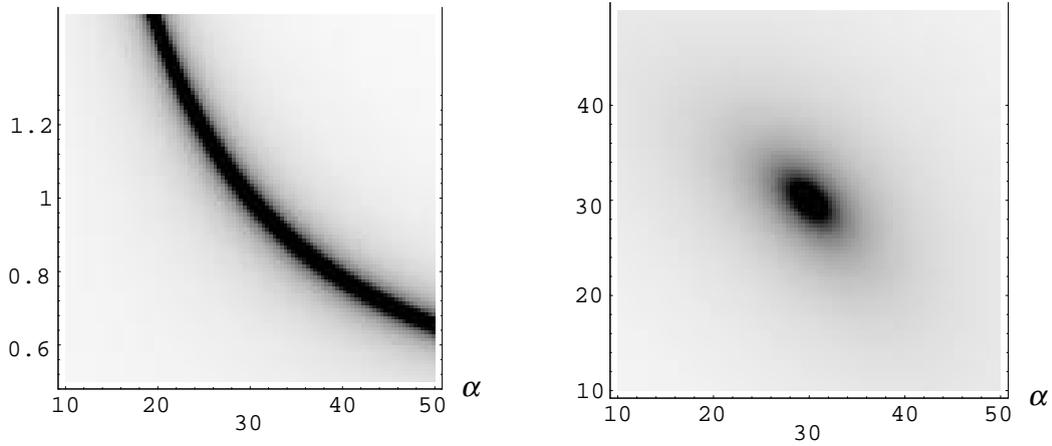


Fig. 17: Two 2D sections through the 4D parameter space $(\alpha, \beta, q_{\text{MK}}, q_{\text{CK}})$ showing their intersections with the singular manifold of the $(0,1,-1,0,1,0)$ -moiré and the forbidden zone surrounding it. (a) A section along the diagonal plane $\alpha = \beta$, $q_{\text{MK}} = q_{\text{CK}}$. (b) A section along the (α, β) plane with $q_{\text{MK}} = q_{\text{CK}} = 1$. Darker shades represent higher moiré periods. Note that the singular point $\alpha = \beta = 30^\circ$, $q_{\text{MK}} = q_{\text{CK}} = 1$ is included in this singular manifold.

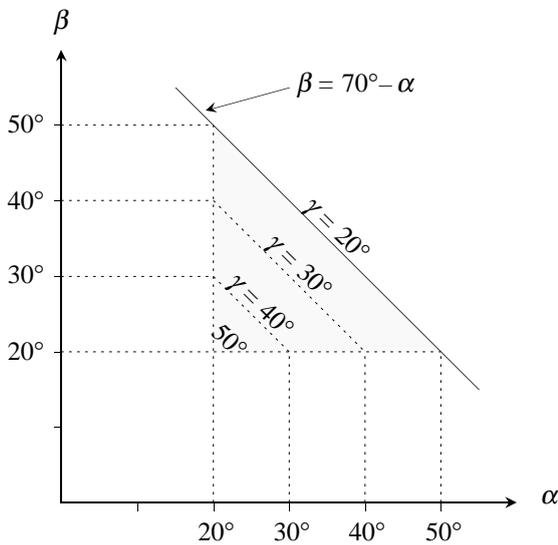


Fig. 18: The admitted range for the angles α , β and γ . Any point outside the gray range violates at least one of the conditions $\alpha \geq 20^\circ$, $\beta \geq 20^\circ$, $\gamma = 90^\circ - (\alpha + \beta) \geq 20^\circ$.

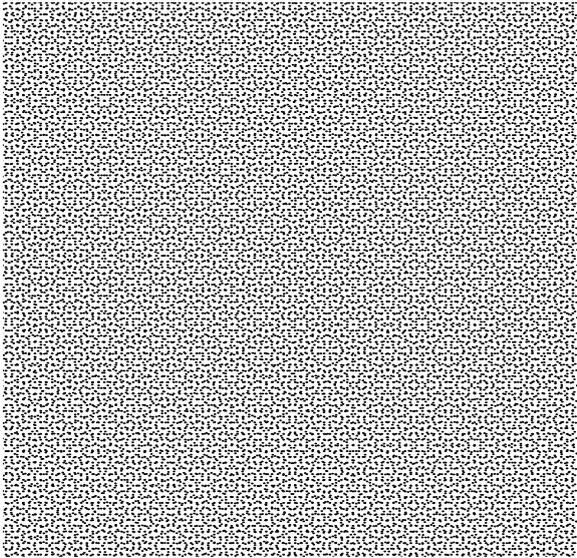


Fig. 19: An example of a stable moiré-free combination of three screens obtained by the algorithm. Small, irregular rosettes, of about the same size as in the conventional 30° - 30° - 30° singular (unstable) moiré-free case, can be seen in the superposition, but no significant moirés (such as in Fig. 9) are present.