Appl. Phys. B 54, 380–395 (1992)

Applied Physics B Photophysics and Laser Chemistry

New Theoretical and Experimental Results in Fresnel Optics with Applications to Matter-Wave and X-Ray Interferometry*

J. F. Clauser and M. W. Reinsch

Physics Department, University of California, Berkeley, CA 94720, USA

Received 3 October 1991/Accepted 17 January 1992

Abstract. We present new methods and formulae for calculating the image amplitude and image spatial power spectral density produced by monochromatic point-source illumination of a finite (and/or infinite) periodic complex transmission grating. At specific finite-width *resonances* the image amplitude is seen to display periodic complex amplitude self-imaging of the grating, with interlaced alias images. The finite width grating resonances (as a function of spatial frequency) are broadened (from zero width) and displaced in frequency relative to those produced by an infinite grating, although the finite resonance width relative to illumination wavelength variation persists with infinite gratings. In the Fresnel domain the self images are generalizations of the Talbot and von Lau effects, while in the Fraunhofer to Fresnel transition domain, our formulae demonstrate the formation of these structures from Fraunhofer diffraction order side-lobes. Using these results, design criteria are provided for constructing lens-free three-grating interferometers with spatially diffuse illumination and detection. Such interferometers have a wide variety of applications for both X-rays and matter-waves, including a phase sensitive imaging device and/or narrow-band interference filter. For wavelengths in the Ångstrom to sub-Ångstrom range they feature high throughput and ease of fabrication. Experimental results using light with such an interferometer are presented. Our results conclusively demonstrate interference and image aliasing in such a device with spatially diffuse illumination and detection. The experiment is readily reproducible in any undergraduate physics laboratory.

PACS: 07.60.Ly, 42.10.Hc, 42.80.Bi

The calculation of a diffraction pattern produced by a transmitting object is a standard problem in physical optics and quantum mechanics. The first step in the calculation is to reduce the governing wave equation (e.g., Maxwell's equations, Schrödinger's equation, etc.) to the Helmholtz equation by factoring out the time dependence. A solution to the Helmholtz equation is then given by the Huygens-Fresnel-Kirchhoff diffraction integral, which allows the image to be calculated in terms of a superposition integral involving the complex grating transmission function and the Green's function for the propagation, sometimes called a propagator or kernel.

When the incoming and outgoing wavefronts have negligible curvature, the Fraunhofer (first order) approximation simplifies the integration, and the image is given by the Fourier transform of the object's complex transmission function, evaluated at spatial frequencies proportional to the lateral position in the image plane [1, 2]. When wavefront curvature cannot be neglected, the Fresnel (second order) approximation is usually required. However, its use introduces quadratic phase factors which typically yield intractable integrals. These may be evaluated numerically. For binary transmission functions associated with an opaque object with one or more apertures, a common procedure is to first calculate the amplitude produced by each edge of the object via Fresnel integrals which are readily evaluated numerically. Next, the amplitude of the whole object is calculated by summing the integrals associated with all of the edges. Unfortunately, analytic approximations for the Fresnel integrals are available only in asymptotic domains. Furthermore, this procedure fails for objects with complex and/or non-binary transmission functions, and a direct numerical integration of the diffraction integral is usually required [3].

Despite the inherent complications associated with quadratic phase factors, when the object is periodic, surprising yet simple periodicities are manifest in the resulting image [17]. In some situations the periodicities give rise to

* Work supported by ONR Grant N00014-90-J-1475, by the firm J. F. Clauser & Assoc., and (MWR) by a U.S. Dept. of Education Fellowship

self-imaging effects for gratings. That is, a grating will produce an exact image of itself at specific image distances. A pair of spaced similar gratings thus produce Moiré patterns (and/or patterns intermediate between Moiré and sinusoidal) in a manner that intimately involves wave interference. Two such self-imaging effects were discovered experimentally by Talbot and von Lau. The scaling of the Talbot effect was first explained by Rayleigh. These effects (discussed in Appendix A) are intimately associated with non-vanishing wavefront curvature and are not produced in the Fraunhofer limit. Indeed, Jahns and Lohmann [4] have pointed out that the Talbot and von Lau effects are reciprocals of each other. They also give a simple geometric construction that allows the effect to be understood as the action of a generalized quadratic Fresnel zone plate in the von Lau geometry with infinitesimal slit widths.

The intractability of the integrals has been attacked in various ways using a variety of mathematical techniques to elucidate (in ways that numerical technique can not) the regularities of the image produced by a periodic structure [5]. Cowley and Moodie [6] used a Fourier series technique to treat the problem of Fresnel diffraction by an infinite breadth periodic amplitude transmission object. Their work thus explains some of the basic periodicities (and 180° phase shifts) of the Talbot and von Lau effects. Appendix B summarizes their result.

It is noteworthy that the Cowley and Moodie treatment is problematic in that it only yields results for objects with infinite extent, but in this limit the second order Fresnel approximation, itself, fails. Unfortunately, their derivation is not readily generalized to the treatment of finite breadth gratings. More important, although additional naturally occurring periodicities are manifest in the Cowley and Moodie formulae, these periodicities are not self evident, and indeed their existence was missed by them [7]. Gori [8] and Sudol and Thompson [9] identified additional image periodicities in the von Lau effect. Sudol and Thompson showed the existence of the intermediate images in the von Lau geometry by calculating the intensity distribution resulting from an infinite grating. Gori showed that intermediate images will exist for a grating composed of a finite number of slits, experimentally demonstrated this effect, and presented qualitative arguments to show that the finite extent of the gratings limited the multiplicity of the intermediate images.

Despite the work listed above, no general analytic method is at hand for determining the image amplitude, along with its periodicities and other interesting features that result from Fresnel diffraction by a finite periodic complex structure. As a result, the detailed structure of and reasons for the intermediate images is obscure. It is the purpose of this paper to present such analytic methods, and further, to point out some of the amazing properties and rich spectroscopy that is latent in a rather innocent looking integral. In Sect. 1 we provide a whole family of equivalent formulae, each of which provides the exact image amplitude (at least to the limits set by the Fresnel approximation and the Huygens-Fresnel-Kirchhoff integral). In Appendix C, we use a method similar to that of Sect. 1 to provide a generalization of Jahns and Lohman's zone plate discussion to the Cowley and Moodie geometry.

In the remainder of Sect. 1 we show that for an increasingly large finite number of grating periods, each having a complex amplitude transmission function, the image *amplitude* for both the primary (Talbot-von Lau) images and the intermediate (aliased) images become exact replicas of a single period's complex amplitude transmission function, except for an extraneous quadratic phase factor. For infinite gratings, our resulting formulae give results in agreement with those derived earlier (reviewed in Appendix B) by Cowley and Moodie. For a modest number of grating periods, the individual image periods become rounded and are limited by a finite extent envelope. Our formulae also elucidate the wiggles seen on images of diffraction patterns observed at the transition between the Fraunhofer and Fresnel regimes. In Sect. 2 we investigate the spectral properties of the image intensity and find a regular spectrum of finite width "resonances" for both finite and infinite numbers of periods, and display the spectrum in closed form for the case wherein each period has the form of a simple rectangular slit.

Since the dimensions associated with grating self-imaging are rarely interesting for applications in the optical domain, the subject of Talbot and von Lau self-imaging is almost never discussed in physical optics textbooks [18]. On the other hand, Fresnel diffraction by transmission gratings assumes great importance at short wavelengths, such as those encountered in matter-wave and X-ray interferometry. As an application of our results that is useful to these arts, in Sect. 3 we describe lens-free grating interferometers that employ a diffuse (spatially incoherent) monochromatic source (uncollimated sodium discharge lamp), and diffuse detector, calculate their spectral properties, and experimentally confirm their operation. Such an interferometer is readily reproduced by a few hours effort in any undergraduate optics laboratory.

1 Amplitude for Fresnel Diffraction by a Finite Periodic Structure

Consider the two dimensional geometry depicted in Fig. 1. A point source at $x = x_0$, $z = z_0$ emits monochromatic waves with unit amplitude at unit distance from it. The waves are passed by a planar, finite extent, infinitesimally thick, periodic transmitting object (grating) located in the $z = z_1$ plane, with its axis of symmetry located on the z-axis. We seek the amplitude $\psi(x_2)$ of the waves on the $z = z_2$ plane at the point (x_2, z_2). Let $R_1 = z_1 - z_0$ be the spacing between the source and grating plane, and let $R_2 = z_2 - z_1$ be the spacing between the grating plane and detector plane. The path length between the source point at (x_0, z_0) and the point $x_1, z_1)$ on the grating plane is given by

$$S_1 = \sqrt{R_1^2 + (x_1 - x_0)^2} \approx R_1 \left[1 + \frac{(x_1 - x_0)^2}{2R_1^2}\right], \qquad (1)$$

where the second expression results from applying the Fresnel approximation. Similarly, the path length between the point at (x_1, z_1) and the point at (x^2, z_2) is given by

$$S_2 = \sqrt{R_2^2 + (x_2 - x_1)^2} \approx R_2 \left[1 + \frac{(x_2 - x_1)^2}{2R_2^2}\right]. \qquad (2)$$

Let us now define the quantity Δ as the path length increase that is due to the lateral displacements $(x_1 - x_0)$ and $(x_2 - x_0)$. It is given by

$$\Delta \equiv S_1 + S_2 - R_1 - R_2\,. \qquad (3)$$

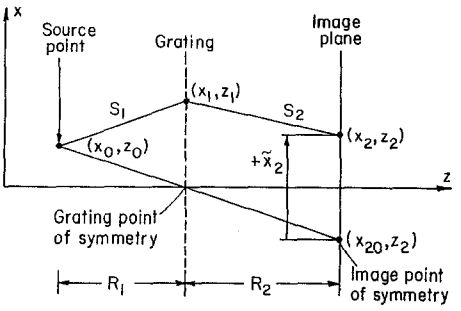

Fig. 1. Layout of the two-dimensional geometry

Combining (1)–(3) and completing the square of the resulting expression, we can write this as

$$\Delta = \Delta_0(x_2, x_0) + \frac{1}{2\varrho}\left(x_1 - \frac{x_2 - x_{20}}{M}\right)^2, \tag{4}$$

where we have defined the reduced radius ϱ as

$$\varrho \equiv \frac{R_1 R_2}{R_1 + R_2},$$

the magnification M as

$$M \equiv \frac{R_1 + R_2}{R_1} = \frac{R_2}{\varrho},$$

the lateral displacement x_{20} of a point on the z_2 plane symmetrically opposite the source point (through grating center) which then is the image centroid as

$$x_{20} \equiv -\frac{R_2}{R_1} x_0,$$

the lateral displacement of an image point $\tilde{x}_2$ from its centroid as

$$\tilde{x}_2 \equiv x_2 - x_{20},$$

and the residual path length $\Delta_0(x_2, x_0)$, that does not depend on the integration variable x_1, as

$$\Delta_0(x_2, x_0) \equiv \frac{(x_2 - x_0)^2}{2(R_1 + R_2)}.$$

If the complex transmission function of the object on the z_1 plane is defined as $t(x_1)$, then the amplitude $\psi(\tilde{x}_2)$ of the wave at the point (x_2, z_2) is given via the Huygens-Fresnel-Kirchhoff diffraction integral [19]. For scalar waves as

$$\begin{aligned}\tilde{\psi}(\tilde{x}_2) &\equiv \psi(x_2) \\ &= \frac{1}{(\mathrm{i}\lambda R_1 R_2)^{1/2}} \int_{-\infty}^{\infty} dx_1 t(x_1) \exp\left(\mathrm{i}\frac{2\pi}{\lambda}\Delta\right) \\ &= \frac{\mathrm{e}^{\mathrm{i}\phi_0}}{(\lambda R_1 R_2)^{1/2}} \int_{-\infty}^{\infty} dx_1 t(x_1) \\ &\quad \times \exp\left[\frac{\mathrm{i}\pi}{\lambda\varrho}\left(x_1 - \frac{\tilde{x}_2}{M}\right)^2\right],\end{aligned} \tag{5}$$

where we further define the extraneous phase factor as

$$\phi_0 \equiv \frac{2\pi}{\lambda}\Delta_0 - \frac{\pi}{4}$$

and shift the origin of the image amplitude to the pattern's centroid. (For a source point on the z-axis and the grating centered on this axis, then $\tilde{\psi} = \psi$). The exonential factor under the integral sign is the Fresnel Green's function.

For large gratings (5) is well known to yield geometric shadows in the small $\lambda\varrho$-limit and Fraunhofer diffraction in the large $\lambda\varrho$-limit. Our purpose is to evaluate it at intermediate wavelengths for finite periodic gratings. We thus prescribe $t(x_1)$ to represent a finite periodic structure, centered on the z-axis. We denote its period as the distance a and its complex transmission for a single period as $s(x_1)$, acting only on the interval $-a/2 \leq x_1 < a/2$, with $s(x_1) = 0$ holding outside of this interval. Suppose that the grating consists of N periods. The grating transmission function $t(x_1)$ is then given by [10]

$$t(x_1) = \sum_{j=(1-N)/2}^{(N-1)/2} s(x_1 - ja). \tag{6}$$

We now use the sifting property of δ-functions

$$s(\xi) = \int_{-\infty}^{\infty} d\xi' s(\xi')\delta(\xi - \xi'), \tag{7}$$

to write the grating transmission function as

$$t(x_1) = \sum_{j=(1-N)/2}^{(N-1)/2} \int_{-\infty}^{\infty} d\xi' s(\xi')\delta(x_1 - ja - \xi'). \tag{8}$$

Substituting (8) into the expression of (5) for $\psi(x_2)$, we have

$$\begin{aligned}\tilde{\psi}(\tilde{x}_2) = \frac{\mathrm{e}^{\mathrm{i}\phi_0}}{(\lambda R_1 R_2)^{1/2}} &\int_{-\infty}^{\infty} dx_1 \sum_{j=(1-N)/2}^{(N-1)/2} \int_{-\infty}^{\infty} d\xi' s(\xi') \\ &\times \delta(x_1 - ja - \xi')\exp\left[\frac{\mathrm{i}\pi}{\lambda\varrho}\left(x_1 - \frac{\tilde{x}_2}{M}\right)^2\right].\end{aligned} \tag{9}$$

Reversing the order of integration, performing the integration over x_1, and using the enforcement by the delta function that

$$x_1 = ja + \xi' \tag{10}$$

holds, we have for the amplitude at (x_2, z_2),

$$\begin{aligned}\tilde{\psi}(\tilde{x}_2) = \frac{\mathrm{e}^{\mathrm{i}\phi_0}}{(\lambda R_1 R_2)^{1/2}} &\int_{-\infty}^{\infty} d\xi' \sum_{j=(1-N)/2}^{(N-1)/2} s(\xi') \\ &\times \exp\left[\frac{\mathrm{i}\pi}{\lambda\varrho}\left(ja + \xi' - \frac{\tilde{x}_2}{M}\right)^2\right].\end{aligned} \tag{11}$$

We now introduce the dimensionless length u by the relation

$$au \equiv \xi' - \frac{\tilde{x}_2}{M}, \tag{12}$$

whereupon (11) becomes

$$\begin{aligned}\tilde{\psi}(\tilde{x}_2) = \frac{a\mathrm{e}^{\mathrm{i}\phi_0}}{(\lambda R_1 R_2)^{1/2}} &\int_{-\infty}^{\infty} du\, s\left(au + \frac{\tilde{x}_2}{M}\right) \\ &\times \sum_{j=(1-N)/2}^{(N-1)/2} \exp\left[\mathrm{i}\pi\frac{\lambda_{\mathrm{TR}}}{\lambda}(u + j)^2\right],\end{aligned} \tag{13}$$

wherein we have introduced the Talbot-Rayleigh wavelength (Appendix A motivates our choice of name) defined as

$$\lambda_{\mathrm{TR}} \equiv a^2/\varrho\,. \tag{14}$$

The single summation in (13) may be converted to two interleaved summations by the self evident formula

$$\sum_{j=(1-N)/2}^{(N-1/2} f(j) = \sum_{q=(1-n)/2}^{(n-1)/2} \sum_{p=(1-r)/2}^{(r-1)/2} f(np+q)\,, \tag{15}$$

where the positive integer, n is a divisor of N, and where the integer r is the ratio

$$r = \frac{N}{n}\,. \tag{16}$$

Using (15), the summation in (13) may then be written as

$$\sum_{j=(1-N)/2}^{(N-1/2} \exp\left[\ \ \right] = \sum_{q=(1-n)/2}^{(n-1)/2} \exp\left[\mathrm{i}\pi\frac{\lambda_{\mathrm{TR}}}{\lambda}(u+q)^2\right] \times \sum_{p=(1-r)/2}^{(r-1)/2} \exp\left(\mathrm{i}\pi\frac{\lambda_{\mathrm{TR}}}{\lambda}n^2p^2\right) \times \exp\left\{\mathrm{i}2\pi\frac{\lambda_{\mathrm{TR}}}{\lambda}[np(u+q)]\right\}. \tag{17}$$

Let us now define the real number μ as

$$\mu \equiv n\frac{\lambda_{\mathrm{TR}}}{\lambda}\,, \tag{18}$$

or equivalently, via the definition of (14),

$$\lambda\varrho \equiv a^2\frac{n}{\mu}\,. \tag{18a}$$

We will eventually focus our attention on cases where μ is approximately an integer. Equivalently, we will focus on cases wherein $\lambda_{\mathrm{TR}}/\lambda$ is approximately equal to a ratio of positive integers m and n [20], as per

$$\frac{\lambda_{\mathrm{TR}}}{\lambda} = \frac{m}{n} + \varepsilon = \frac{\mu}{n}\,, \tag{19}$$

where the integer n, introduced above, is a divisor of N and, where the absolute value of the remainder ε is small (see Sect. 1.3). Using (19), the sum over j of (17) may be expressed as

$$\sum_{j=(1-N)/2}^{(N-1)/2} \exp\left[\ \ \right] = \sum_{q=(1-n)/2}^{(n-1)/2} \exp\left[\mathrm{i}\pi\frac{\mu}{n}(u+q)^2\right] \times \sum_{p=(1-r)/2}^{(r-1)/2} \exp\left(\mathrm{i}2\pi p\frac{mn}{2}\right) \times \exp(\mathrm{i}\pi\varepsilon n^2p^2) \times \exp[\mathrm{i}2\pi\mu p(u+q)]\,, \tag{20}$$

where, in reducing this expression, we have used the identity $\exp(\mathrm{i}\pi mnp^2) = \exp(\mathrm{i}\pi mnp)$, that is true because the products (mnp) and (mnp^2) are both even or odd integers simultaneously.

Collecting terms, we have for the amplitude at the point (x_2, z_2)

$$\tilde{\psi}(\tilde{x}_2) \approx \frac{a\mathrm{e}^{\mathrm{i}\phi_0}}{(\lambda R_1R_2)^{1/2}} \int_{-\infty}^{\infty} du\, s\left(au + \frac{\tilde{x}_2}{M}\right) \times \sum_{q=(1-n)/2}^{(n-1)/2} \exp\left[\mathrm{i}\pi\frac{\mu}{n}(u+q)^2\right] \times \sum_{p=(1-r)/2}^{(r-1)/2} \exp\left(\mathrm{i}2\pi p\frac{mn}{2}\right) \exp(\mathrm{i}\pi\varepsilon n^2p^2) \times \exp[\mathrm{i}2\pi\mu p(u+q)]\,. \tag{21}$$

We can simplify this expression by moving the summation to include the factor s, by changing the variable of integration to

$$v = -\mu(u+q) \tag{22}$$

and by using the definitions of ϱ and λ_{TR}, whereupon (21) may be written as

$$\tilde{\psi}(\tilde{x}_2) \approx \frac{(n/\mu)^{1/2}\mathrm{e}^{\mathrm{i}\phi_0}}{(R_1+R_2)^{1/2}} \int_{-\infty}^{\infty} dv\, G(v)\, t'\left(\frac{\tilde{x}_2}{M} - \frac{av}{\mu}\right), \tag{23}$$

for which we have defined a modified n-period grating transmission function t' as

$$t'(x) \equiv \frac{1}{n}\sum_{q=(1-n)/2}^{(n-1)/2} s(x-aq)\,, \tag{24}$$

and further introduce the dimensionless Green's function $G(v)$ as a product of two functions

$$G(v) = E(v)P\left(v - \frac{mn}{2}\right). \tag{25}$$

The first factor, $E(v)$, is an aperiodic function defined as

$$E(v) = \exp\left(\frac{\mathrm{i}\pi}{\mu n}v^2\right), \tag{26}$$

which we shall refer to as the *envelope factor*. The second factor, $P(v)$, will be recognized as a Fourier series, and thus, as a periodic function with unit period for r odd and period 2 (with a symmetrically inverted second half period) for r even. It is defined as

$$P(v) = \sum_{p=(1-r)/2}^{(r-1)/2} \exp(\mathrm{i}\pi\varepsilon n^2p^2)\exp(-\mathrm{i}2\pi pv)\,. \tag{27}$$

It is noteworthy, however, that even though the function $P(v)$ is periodic, the function $G(v)$ is not. For a given choice of n, (23)–(27) yield an exact result (at least to the limits set by the Fresnel approximation and the Huygens-Fresnel-Kirchhoff integral).

Our formulation of the Fresnel diffraction problem provides a solution, (23), in a particularly simple form that exhibits important features that were previously obscure. The image's periodicities are provided by the functions t and G, while the envelope is determined by the function E. In a parallel to the usual result the image is calculated in terms of a

superposition integral involving the complex grating transmission function and a Green's function for Fresnel propagation. Indeed, for the case $n = N$, the formulae are essentially the same. However, for the $\mathcal{N} - 1$ cases for $n \neq N$ (where $\mathcal{N}$ is the number of distinct divisors of N) our $\mathcal{N}-1$ forms are notably different from the $n = N$ result. The superposition integral now involves the Green's function $G(v)$, and a *shortened* n-period grating $t'(x)$ (with the argument exhibiting the geometric magnification). Moreover, periodicity is shifted from the grating function to the function $P(v)$. When the various grating period self-images do not overlap and the phase advance rate that is due to the envelope factor $E(v)$ is slow, the phase introduced by the factor $E(v)$ causes no interference and the intensity displays the periodicity associated with the periodic function $P(v)$. It is this change in structure that now allows image periodicities to be self-evident.

1.1 Choice of n and Form of P(v) for Small $|\varepsilon|$

A natural question to ask is which of the $\mathcal{N}$ values for n should we choose, since all choices lead to valid formulations. From the numerical point of view the choice may be only a matter of taste or expedience. However, the motivation for formulating the result as we have done is that when the choice is made so that μ is nearly an integer m, then the formulae reduce to a particularly simple form. Let us thus consider the periodic function, $P(v)$, given by (27) and examine situations wherein $\mu \approx m$ (and thus $\varepsilon \approx 0$) holds. For small $|\varepsilon|$, the factor $\exp(\mathrm{i}\pi\varepsilon n^2 p^2)$, effects only the terms in the summation with large $|p|$, which, in turn, is limited by the sum to the domain $|p| \leq r/2$. Thus, whenever the inequality

$$|\varepsilon| \leq \frac{2}{r^2 n^2} = \frac{2}{N^2} \tag{28}$$

holds (readily achieved for a modest number of periods), the dephasing of the various terms of the summation by this factor is less than $\pi/2$, whence the factor may be omitted. Thus, when (28) is satisfied, we may approximate $P(v)$ as

$$P(v) \approx \sum_{p=(1-r)/2}^{(r-1)/2} \exp(-\mathrm{i}2\pi p v)\,. \tag{29}$$

Equation (29) represents the sum of a geometric series that evaluates as

$$P(v) \approx \frac{\sin(\pi r v)}{\sin(\pi v)}\,. \tag{30}$$

Since $r = N/n$ is an integer, we recognize $P(v)$ in the present form as the periodic interference function that results from Fraunhofer diffraction by a periodic structure [2]. It has the form of an infinite series of spikes with period one or two. When r is an odd integer, then all spikes are positive and the function has period one; when it is even, the spikes have alternating sign and the function has period two. When r is large (i.e., when the number of grating periods N is large), then each spike approaches a δ-function, and the function approaches a comb-function (or sum of two laterally displaced opposite sign comb-functions). These near δ-functions then give rise to grating amplitude self-imaging.

The choice $n = N$, $r = 1$ yields a constant for $P(v)$ and, correspondingly, yields an expression identical to our original form (5), providing nothing new. For the form of (30) to be useful we desire a small value for n to be chosen, or equivalently, a value for $r = N/n \gg 1$. A second question thus arises concerning the uniqueness of any given choice for n. Suppose we have two choices n and n', thereby defining m and m' and remainders ε and ε' by (19) as

$$\frac{\lambda_{\mathrm{TR}}}{\lambda} = \frac{m}{n} + \varepsilon = \frac{m'}{n'} + \varepsilon'\,, \tag{31}$$

where m and m' are integers chosen to respectively minimize ε and ε'. By the triangle inequality and (28) we have

$$\left|\frac{m}{n} - \frac{m'}{n'}\right| = |\varepsilon - \varepsilon'| < \frac{2}{N^2}\,. \tag{32}$$

Multiplying by nn' (32) can be rewritten as

$$|mn' - m'n| < \frac{2n}{N}\frac{2n'}{N} \ll 1\,, \tag{33}$$

where the last inequality is due to our restriction $n, n' \ll N$. But the quantity $|mn' - m'n|$ is a non-negative integer and thus must be zero. Hence we have $m/n = m'/n'$ and the choice for n is thus unique, except to a common factor of m and n.

1.2 Periodic Grating Self-Imaging for Finite N

To get a feeling for the solution, consider a particularly simple case in the limit as $N \to \infty$ and $\varepsilon \to 0$, with $m = 1$, $n = 1$, $r = N$, and N odd. In such case the function $P\left(v - \frac{1}{2}\right)$ becomes an infinite periodic string of δ-functions on the half-integers, and the function E is a constant, $E = i$. The function $t'(x)$ has only one period and reduces to $s(x)$. The superposition integral then produces an image that is a periodic string of magnified single periods, all with the same phase (except for an extraneous quadratic phase factor), i.e., an exact magnified image of the original infinite grating. The image is offset, however, from the position of the geometric shadow of the grating by one half period because of the term $mn/2 = 1/2$ in the argument of P. The image is thus periodic with period Ma, multiplied by the extraneous quadratic phase factor $\exp[\mathrm{i}\phi_0(x_2, x_0)]$.

Let us examine the argument of the n-period grating function $t'\left(\frac{x_2}{M} - \frac{av}{\mu}\right)$. When N is finite and ε is still sufficiently small that (28) holds, then the central part of the image is given by a convolution of this function with a string of spikes that is periodic with unit period, except for a different phase of each spike. The convolution of each spike of $P\left(v - \frac{mn}{2}\right)$ with each period of the function t' produces a geometrically magnified image of a single grating period. The n periods of the function t' produce n such images. When $\mu = m$ holds then some of the various images may lie exactly on top of each other. Depending on the phase they may either cancel or reinforce each other. In general, when the fraction m/n is not reduced to lowest terms, destructive interference occurs. When the single period func-

tion is negligible everywhere except for an isolated feature of width $s < (a/\mu)$ (as in the case of a grating composed of narrow slits) then the contributions from neighboring spikes will not overlap and the phase variation from one spike to the next will be inconsequential. In such case the integration produces an image that is (except for phase and envelope) a periodic string of magnified single period images.

The image period is Ma/μ. That is, the image has μ periods per geometric shadow period Ma. We thus call the quantity $\mu \approx m$ the *alias multiplicity*. We note that the Cowley and Moodie result (Appendix B, (B6)) indicates an image whose period is Ma. When $\mu = m$ is an integer, then the period Ma/m is commensurate with the period Ma and is compatible with the Fourier series of (B6). However, when μ is not an integer, then the periods Ma/μ and Ma are incommensurate and our solution provides a frequency shift from that of Cowley and Moodie. It is noteworthy that a new feature arises here with finite gratings of modest extent that was not present in the infinite grating case treated by Cowley and Moodie. (Spectral properties of the image will be discussed in Sect. 2).

Next, consider a second simple case with $m = 2$, $n = 1$, $r = N$, and N odd, in the limit as $N \to \infty$ and $\varepsilon \to 0$. Since the term $mn/2$ is now integral, the lateral half-period shift that was due to this term vanishes. However, there will be images of the original magnified single period both in the position of the geometric shadow as well as aliased images half-way in between the shadows.

Let us examine further the half-peiod shift. To do so, consider another illustrative simple case with $m = 1$, $n = 3$, $r = N/3$, and N even, again in the limit $N \to \infty$ and $\varepsilon \to 0$. Since $mn/2$ is half integral, there will be a half-period shift via the argument of P. The grating t' now has three periods, spaced in v at integral values of $\tilde{x}_2/(Ma)$, i.e., spaced with the same interval as the period of the function P. Thus, the three images produced by each of the three periods of the shortened grating function t' will coincide. However, these three periods occur on integral values of v, while the geometric shadows of the original grating occur on half-integral values of $\tilde{x}_2/(Ma)$. A Second half-period lateral shift of the image is thus introduced. The two lateral shifts will cancel and the images will lie on the positions of the geometric shadows.

We have noted two sources of half-period shifts of the image. When the product mn is odd, the spikes occur at v equal to half integer values, while when it is even, they occur at integer values. Whenever n is odd but r is even, the positions of the grating periods in the function t will be offset from those of the original grating function t by an additional half-period. Otherwise, the positions of the grating periods in the function t' will coincide with those of the original grating function t. In some cases, the shifts cancel. The various possibilities are summarized in Table 1.

1.3 Form of $P(v)$ when $|\varepsilon|$ is not Small

When the number of periods N is large, then (28) sets rather stringent limits on the variation of $|\varepsilon|$ and demands a careful matching of λ_{TR}/λ to m/n for the function $P(v)$ to take the form given by (30), which, in turn, has the form of a periodic series of narrow spikes, each of which approaches a δ-function in the $N \to \infty$ limit. However, the function $P(v)$ displays remarkable properties. Even when (28) is strongly violated, the function $P(v)$ in a convolution still acts with the same effect. Thus, (28) does not represent a requirement for self-imaging to occur.

Figure 2a–d show a numerical evaluation of the modulus and phase of one period of $P(v)$ as a function of v for various values of ε for the case $n = 1$ and $r = N = 75$. As anticipated above, when (28) holds, each period has the form of a real spike. For small $|\varepsilon|$ such that (28) holds and the full width of each spike, from (30), is given by

$$\Delta v_{\text{spike}} = \frac{2n}{N} = \frac{2}{r}. \tag{34}$$

However, as $|\varepsilon|$ increases above the limit set by (28), the function becomes complex. The modulus of each spike becomes a narrow square topped plateu. The phase is stationary at the plateau's center and increases quadratically outward. The full width of the plateau is approximately given by

$$\Delta v_{\text{plateau}} \approx Nn|\varepsilon| + \frac{2n}{N}. \tag{35}$$

(Actually, this Formula will slightly underestimate the plateau width, depending upon how the width is defined.) By considering the form of the Fresnel integrals, we note that in the convolution with a slowly varying function the effective spike width will be determined not by the plateau width, but by the quadratic phase factor, which then provides a narrower effective spike width. Indeed, a self-image cannot display features of a single period (such as a slit of width s) narrower than a Δv_{phase} given approximately by

$$\Delta v_{\text{phase}} \approx \sqrt{3|\varepsilon|} + \frac{2n}{N}. \tag{36}$$

Table 1.

N	n	r	m	mn	j (6)	q (15)	Offset 1/2 period		
							P	t'	Image
Odd	Odd	Odd	Even	Even	Int	Int	No	No	No
Odd	Odd	Odd	Odd	Odd	Int	Int	Yes	No	Yes
Even	Even	Odd	Odd	Even	1/2 Int	1/2 Int	No	No	No
Even	Even	Odd	Even	Even	1/2 Int	1/2 Int	No	No	No
Even	Even	Even	Odd	Even	1/2 Int	1/2 Int	No	No	No
Even	Even	Even	Even	Even	1/2 Int	1/2 Int	No	No	No[a]
Even	Odd	Even	Odd	Odd	1/2 Int	Int	Yes	Yes	No
Even	Odd	Even	Even	Even	1/2 Int	Int	No	Yes	Yes[b]

[a] This case does not occur in practice since the ratio m/n is not reduced to lowest terms, and as mentioned in the text will cause destructive interference and vanish

[b] This displacement is due solely to displacement of the shadow positions

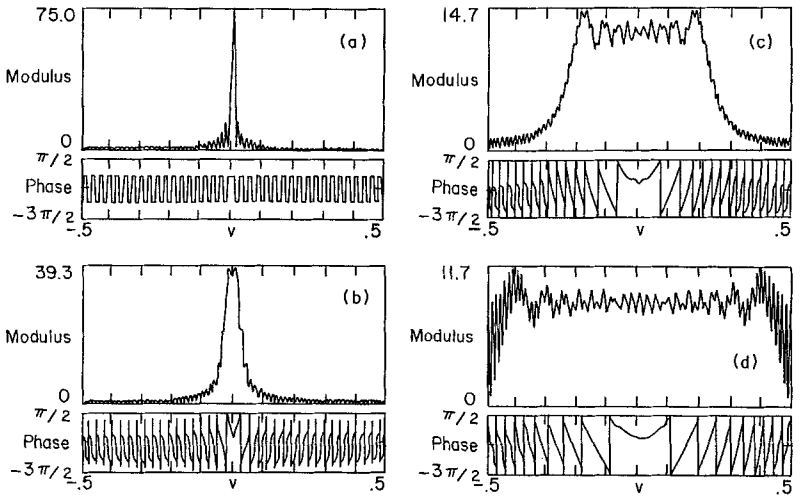

Fig. 2a–d. Numerical evaluation of the modulus and phase of one period of $P(v)$ as a function of v for various values of ε for $n = 1$ and $r = N = 75$. **a** $\varepsilon = 0$, **b** $\varepsilon = 1/(16N)$, **c** $\varepsilon = 1/(2N)$, **d** $\varepsilon = 1/N$

Thus each spike is still effectively narrow as long as the inequality

$$|\varepsilon| \ll \frac{1}{nN} \tag{37}$$

holds. When (37) becomes an equality, then the plateau width has spread to one whole period, and is approximately constant for all v. For larger $|\varepsilon|$ the plateau abruptly disintegrates and becomes oscillatory. For even larger $|\varepsilon|$ such that μ/n again closely approximates a ratio of different integers, (with n now a different divisor of N) then $P(v)$ reforms into a new set of spikes with a new multiplicity appropriate to the new n and nearest m. Although we have no ready proof for assertions (35) and (36), computationally they appear to be approximately true for all the cases we have evaluated. The limit set by (37) may be found by examining the results of Sect. 3 in the limit as $\lambda/a \to 0$.

The area for each period (evidently a spike and/or oscillatory-phase plateau) of $P(v)$ is readily calculated. When r is odd, a term-by-term integration of the sum in (27) from $v = -1/2$ to $v = +1/2$ yields zero for all terms in the sum except for the one with $p = 0$, whose integral is one, independently of the value of r (odd), ε, and n. When r is even, then p takes on half integral values and the integration is more complicated. A term-by-term integration of the sum in (27) from $v = -1$ to $v = +1$ (or from $-1/2$ to $+3/2$) yields zero for all terms, consistent with period 2 alternating sign spikes. For given values of v, r, and ε we may define

$$P_o\left(v, \frac{r-1}{2}, \varepsilon\right) \equiv P(v) \quad \text{for } r \text{ odd}, \quad \frac{r-1}{2} \text{ integral},$$

and similarly

$$P_e\left(v, \frac{r-1}{2}, \varepsilon\right) \equiv P(v),$$

$$\text{for } r \text{ even}, \quad \frac{r-1}{2} \text{ half-integral}.$$

It is then straightforward to show that

$$P_e\left(v, \frac{r-1}{2}, \varepsilon\right) = \frac{1}{2}\left[P_o\left(\frac{v}{2}, r-1, \frac{\varepsilon}{4}\right) - P_o\left(\frac{v-1}{2}, r-1, \frac{\varepsilon}{4}\right)\right] \tag{38}$$

holds, whereupon P_e exhibits alternating-sign spikes and all spikes with the same sign are periodic with period 2. Thus, P_e may be expressed as a difference of two periodic functions, one whose integral from -1 to $+1$ is 1, and one whose integral is -1.

Thus, each spike's area is effectively constant. Since the effective width approaches zero as $N \to \infty$, the periodic function $P(v)$ still acts like a periodic string of delta functions (with unit period) as long as μ differs from an integer by less than $\frac{1}{2N}$. For large but finite N, (37) clearly allows a far wider range of $\varepsilon \neq 0$ over which *near* or *filtered* self-imaging (i.e., with rounded features) occurs than does (28). For infinite N, exact self-imaging occurs only when ε exactly equals zero. For $N \to \infty$ and $\varepsilon \neq 0$, an alternative choice for m and n is thus required, although the resulting image multiplicity may be large. For wide single period features, the effect may result in piled-up (and thus, complicated) self-images.

Thus, with $|\varepsilon|$ within the limit set by (37) (much wider than the limit set by (28)), as N approaches infinity, the function $P(v)$ maintains the form of an infinite-string of constant area spikes whose width approaches zero as $N \to \infty$. In the $N \to \infty$ limit both (28) and (37) require μ to be an exact integer, in order for the function $P(v)$ to become a periodic string of δ-functions. Since the periods Ma and Ma/m are commensurate for $\mu = m$, we note that this result is consistent with the $N \to \infty$ Cowley and Moodie solution, summarized in Appendix B, which yields an image amplitude that is always periodic with period Ma. Thus, self-imaging then occurs for any $\mu = m$ and n for which the multiplicity m is sufficiently small that the self-images do not overlap.

1.4 Self-Image Envelope Width and the Effect of $E(v)$

The infinite string of spikes would create an infinite number of single period images, were it not for the action of the factor $E(v)$ which gives rise to a finite envelope. At large v the factor $E(v)$ will cause the phase of the Green's function to vary more and more rapidly. The phase ϕ shift introduced

by this factor is given by

$$\phi = \frac{\pi v^2}{\mu n}. \tag{39}$$

For a grating displaced from the origin [in the v-space convolution of (23)] to a value $\tilde{v}$ given by

$$\tilde{v} \approx \frac{\tilde{x}_2 \mu}{Ma}, \tag{40}$$

the phase advance rate is then given by

$$\frac{d\phi}{dv} = \frac{2\pi\tilde{v}}{\mu n} = \frac{2\pi\tilde{x}_2}{nMa}. \tag{41}$$

The image amplitude produced by a given spike's contribution to the convolution will diminish at values of $\tilde{v}$ for which the phase change provided by $E(v)$ across its finite width is greater than 2π, or equivalently when

$$2\pi = \Delta v_{\text{spike}} \frac{d\phi}{dv} \tag{42}$$

holds. The factor $E(v)$, thus provides an envelope full-width given by

$$\Delta\tilde{x}_2 \approx MaN \quad \text{and} \quad \Delta\tilde{v} \approx mN \tag{43}$$

which, for $\mu \approx m$, is just the magnified geometric-shadow grating width. Correspondingly, subsidiary maxima are also to be expected when the phase change across a spike is an odd multiple of π.

1.5 Diffraction Regimes and the Fraunhofer Limit

Various diffraction regimes are specified by the wavelength and the geometry. To do so, a few definitions are needed. First, we define the grating effective full-width w_g as

$$w_g \equiv aN. \tag{44}$$

Correspondingly, in v-space we define the shortened grating's full width, $\Delta v_{\text{sh.gr.}}$, which, by (23) and (24), may be written as

$$\Delta v_{\text{sh.gr.}} = n\mu. \tag{45}$$

We further define the dimensionless ratio $\mathscr{F}$ as

$$\mathscr{F} \equiv \frac{w_g^2}{\varrho\lambda} = N^2 \frac{\lambda_{\text{TR}}}{\lambda} = N^2 \frac{\mu}{n}$$
$$= Nr\mu = r^2 n\mu = r^2 \Delta v_{\text{sh.gr.}}, \tag{46}$$

where the various evaluations of $\mathscr{F}$ have used (16), (18a), and (44)–(45). In terms of $\mathscr{F}$ the multiplicity is then given as $\mu = \mathscr{F}/(Nr)$ and the short grating width in v-space is given as $\Delta v_{\text{sh.gr.}} = \mathscr{F}/r^2$.

Fraunhofer diffraction orders always occur at diffraction angles spaced by

$$\theta_{\text{Fraunhofer}} = \ell \frac{\lambda}{a}, \tag{47}$$

where ℓ is an integer. However, the discussion of Sect. 1.2 indicates that generalized Talbot fringes [i.e., peaks in $P(v)$] will occur at angles spaced by

$$\theta_{\text{Talbot}} = \ell \frac{Ma}{\mu R_2} = \ell \frac{\lambda}{na}. \tag{48}$$

Consider our various equivalent formulae for the image amplitude. The smallest r Formula has $r = 1$, at $n = N$, whereupon the function $P(v)$ is no longer periodic but instead constant ($= 1$), and the "spike" width is effectively infinite. This formula is the usual expression for Fresnel diffraction. The minimum value for the integer r, for which $P(v)$ is a periodic function, is 2 (with $n = N/2$). It corresponds to the widest spikes in $P(v)$, since, by (27), this function takes the form of a cosine.

The maximum value for the integer r is the largest divisor of N, i.e., it is $r = N$, itself, with $n = 1$. It, correspondingly, yields $\mu = \mathscr{F}/N^2$ and $\Delta v_{\text{sh.gr.}} = \mathscr{F}/N^2$. By (34) the largest r formula corresponds to the narrowest spike width, and by (45) the narrowest short grating width. For any r the ratio of short grating width to spike width is $\mathscr{F}/(2r)$. Thus, the maximum r formula minimizes this ratio.

Fraunhofer diffraction orders are easily seen to be formed using the maximum r formula, whereas with the usual Fresnel $r = 1$ formula, their presence is less immediately evident. From (47) and (48) we see that for $n = 1$ and $r = N$ the angles $\theta_{\text{Fraunhofer}}$ and θ_{Talbot} are equal. This value of n gives only one period to the short grating. We may now specify "long wavelength" relative to $2w_g^2/(N\varrho)$ as determining the Fraunhofer regime, i.e., this regime now corresponds (for modest N) to values of the parameter $\mathscr{F}$ satisfying equivalent inequalities

$$\mathscr{F} \lesssim \frac{N}{2}, \qquad \frac{\lambda_{\text{TR}}}{\lambda} \equiv \frac{\mu}{n} \lesssim \frac{1}{2N}. \tag{49}$$

In such case the short grating will always fit within the central spike of $P(v)$. For (49) strongly satisfied, the image amplitude will follow the form $P(v)$, with its peaks at the Fraunhofer maxima, broadened but slightly by the finite width of the grating single period. Thus, there is no coincidence in our observation of Sect. 1.1 that the function $P(v)$ reduces to the periodic interference function that results from Fraunhofer diffraction by a periodic structure.

Consider next what happens when (49) is minimally violated. Since the $n = 1$ case single period grating now has dimensions comparable to the spike, the image will be broadened proportionally to the grating's width. For (49) an approximate equality and $\mu \approx 1$ holds, then μ is approximately an integer. As we have seen, generalized Talbot-von Lau fringes will form. Comparing (47) with (48), there will be $n = N$ of these per order spacing. Indeed, the standard Fraunhofer interference function as in [2] predicts a side lobe period equal to the order spacing divided by N. In the near Fraunhofer regime, the fringes will appear as wiggles superimposed on the basic Fraunhofer order structure and will form with increasing $\mathscr{F}$ from remnants of the order side-lobes.

Finally, consider what happens for even shorter λ, such that (49) is more strongly violated. Then, n may take on other integer values less than N, and/or μ may take on integer values greater than one, whereupon n generalized μ, n Talbot fringes will form per Fraunhofer order separation.

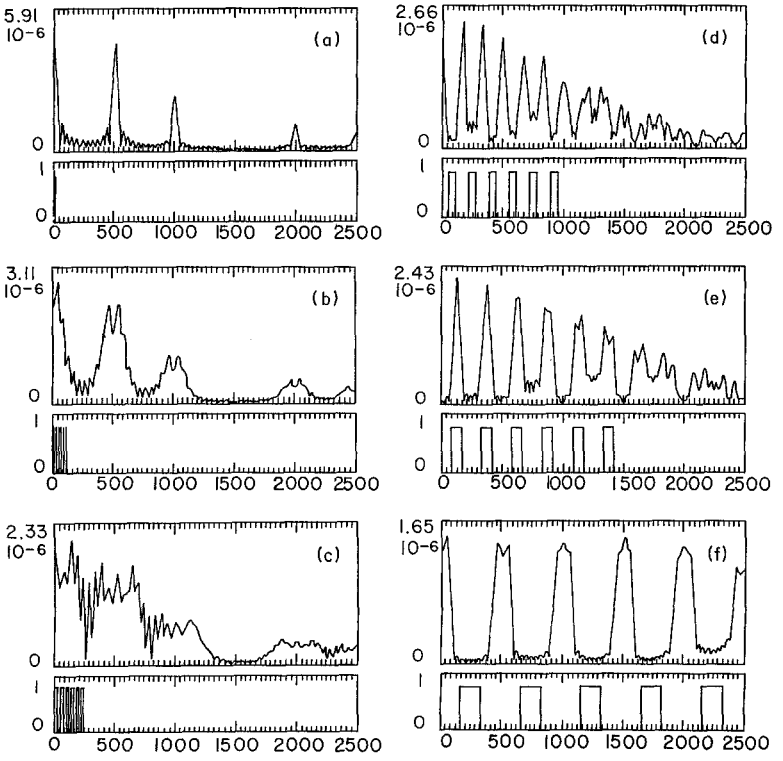

Fig. 3a–f. Image intensity and associated geometric shadow patterns (symmetric about the origin, horizontal-axis $= \tilde{x}_2$, in units $= 10^{-6}$) showing the transition from Fraunhofer diffraction orders to generalized Talbot-von Lau fringes for various R_1 and for a grating with $N = 12$ slits, $\mathscr{s}/a = 1/3$, $R_2 = 1$, $a = 10^{-6}$, and fixed $\lambda = 0.5 \times 10^{-9}$. Fraunhofer orders occur at the major horizontal-axis divisions, i.e., at integral multiples of 500×10^{-6}, while the horizontal-axis subdivisions are at the order spacing N. **a** has $R_1 = \infty$, $\mu/n = 0.002$, and (49) satisfied. **b** has $R_1 = 0.0504$, $\mu/n = 1/24$, and (49) an equality. **c** has $R_1 = 0.0246$, $\mu/n = 1/12$, (49) violated. **d–f** correspond to cases with $\mu = m = 1$, with (52) violated and with $n = 3$ ($R_1 = 0.006$), $n = 2$ ($R_1 = 0.004$), and $n = 1$ ($R_1 = 0.002$), respectively

Finally, consider what happens for even shorter λ, such that (49) is more strongly violated. Then, n may take on other integer values less than N, and/or μ may take on integer values greater than one, whereupon n generalized μ, n Talbot fringes will form per Fraunhofer order separation.

Examples of these features are shown in Fig. 3a–f. Here we have numerically calculated the image intensity using the usual but equivalent $r = 1$ formula, from which none of the above behavior is immediately evident. For all of these figures the calculation is for a grating composed of 12 unit-transmission slits [21] of width $\mathscr{s}$, with $\mathscr{s}/a = 1/3$, $R_2 = 1$, $a = 10^{-6}$, and $\lambda = 50 \times 10^{-9}$ held constant, thereby freezing the positions of the Fraunhofer diffraction orders at integral multiples of 500×10^{-6}. Among these figures we vary the value of R_1 and, correspondingly, vary the value of $\mathscr{F}$. (For this geometric variation all alias images lie on top of each other. If instead we vary the wavelength, keeping the geometry fixed, then the positions of the Fraunhofer shift and aliases are evident.)

Inequality (49) is satisfied in Fig. 3a, b, and well formed diffraction orders are evident. The modulating wiggle period (except at the orders themselves, wherein the associated spike width is twice the wiggle period) is the order spacing divided by n. Figure 3c corresponds to a case ($n = N$, $\mu = 1.0$) wherein (49) is violated by only a factor of two, with only a hint of the order of structure remaining. Figures 3d–f, all correspond to cases with $\mu = m = 1$, with (49) violated, and with $n = 3$, $n = 2$, and $n = 1$, respectively. Figure 3d–f each show n equally spaced well-formed fringes per order separation.

In these calculations the number of slits N is modest. Simpson's rule is employed for the integration using 51 points per open slit portion. Hence, the self-images are rounded. Similar numerical calculations with larger numbers of slits (and/or greater numerical resolution) lead to, correspondingly, greater resolution in the self-images. For gratings composed of simple slits (as in these examples) rectangular fringes with evident Gibbs phenomena are always formed when self-images are expected.

Finally, we note that (49) represents a generalized condition for Fraunhofer diffraction to occur relative to the conditions presented in [2] (Chap. 8.3, (32)–(34), p. 384), which instead require $R_1\lambda \ll w_g^2/4$, and $R_2\lambda \ll w_g^2/4$ to both hold [their Ineqs. (33)]. Physical optics of finite periodic structures now may be divided into the following sequence of four regimes with decreasing wavelength, each with its own rich structure: Fraunhofer, Fraunhofer-to-Fresnel transition, Talbot-von Lau, and geometric shadow.

2 Spatial Power Spectrum for Fresnel Diffraction by a Finite Periodic Structure

A knowledge of the spatial power spectral density in the image plane is useful for a calculation of the net transmitted flux when a Moiré is formed by placing a second periodic transmission grating in the image plane. We thus proceed to calculate the spectral properties of the image intensity spatial distribution. Using (5) and the notation developed above, the image intensity is given by

$$|\tilde{\psi}(\tilde{x}_2)|^2 = \frac{1}{\lambda R_1 R_2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, t(x) t^*(y) \times \exp\left\{\frac{i\pi\mu}{a^2 n}\left[\left(x - \frac{\tilde{x}_2}{M}\right)^2 - \left(y - \frac{\tilde{x}_2}{M}\right)^2\right]\right\}, \quad (50)$$

where the two x_1 integrations use the dummy variables x and y. Upon expanding the squares, canceling terms, and

extracting a common factor within the exponent, the exponential factor may be simplified to

$$\exp\left\{-\frac{i2\pi\mu}{n}\left(\frac{y-x}{a}\right)\left[\left(\frac{y+x}{2a}\right)-\frac{\tilde{x}_2}{Ma}\right]\right\}.$$

The integration is more readily performed in a rotated and rescaled coordinate system, shown in Fig. 4, for which we define the transformation

$$\xi \equiv \frac{y+x}{a}, \quad \eta \equiv \frac{y-x}{a}, \tag{51}$$

with a Jacobian $a^2/2$. In this coordinate system we may write the intensity distribution as

$$|\tilde{\psi}(\tilde{x}_2)|^2 = \frac{a^2}{2\lambda R_1 R_2}\int_{-\infty}^{\infty} d\eta \int_{-\infty}^{\infty} d\xi \times t\left[\frac{a(\xi-\eta)}{2}\right] t^*\left[\frac{a(\xi+\eta)}{2}\right] \times \exp\left[-2\pi i\frac{\mu}{n}\eta\left(\frac{\xi}{2}-\frac{\tilde{x}_2}{Ma}\right)\right]. \tag{52}$$

Factoring the term in the exponential that is independent of ξ, (52) may be written as

$$|\tilde{\psi}(\tilde{x}_2)|^2 = \frac{a^2}{\lambda R_1 R_2} \times \int_{-\infty}^{\infty} d\eta \left[\exp\left(2\pi i\eta\frac{\mu}{n}\frac{\tilde{x}_2}{Ma}\right)\right] F(\eta), \tag{53}$$

where we define $F(\eta)$ as

$$F(\eta) \equiv \frac{1}{2}\int_{-\infty}^{\infty} d\xi\, t\left[\frac{a(\xi-\eta)}{2}\right] t^*\left[\frac{a(\xi+\eta)}{2}\right] \times \exp\left(-\pi i\frac{\mu}{n}\eta\xi\right). \tag{54}$$

For suitably renormalized $\tilde{x}_2$, we recognize (53) as the inverse Fourier transform of $F(\eta)$. Hence, the function $F(\eta)$ represents the spatial power spectrum of the image intensity.

The single slit periods shown in Fig. 4 form a finite raster, periodic with the same period in both x and y. For such a raster the double integral may be represented as a finite double summation over the unit-cell double integral. When the single period function $s(x)$ has a simple form, then both integrals may be evaluated explicitly. For example, when the single period function $s(x)$ has the form corresponding to a slit with a width $\mathscr{s}$ less than half the period a, then the function $F(\eta)$ can be represented as a finite double summation of unit-cell double integrals. (A similar procedure can also be carried out when $\mathscr{s}/a$ is greater than 1/2, however, it is more complicated.)

Viewing the ξ-integral as a function of η in the rotated coordinate system of Fig. 4, we see that the power spectrum F is zero for values of η that slot diagonally between the now diamond-shaped raster elements and nonzero only for values of η that intersect the raster elements. Intersection occurs only for values of η within $\mathscr{s}/a$ of an integer. The raster thus renders the form of the power spectrum to be a

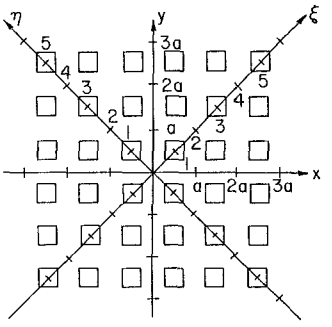

Fig. 4. Rotated and rescaled coordinate system used for performing the integration of the spatial power spectral density. The side of each square is $\mathscr{s}$

set of spikes or interference "resonances". This assertion is true whether or not the number of slits is infinite. For a finite number of slits, there will be a finite number of resonances, and vice versa. For either a finite or an infinite number of slits the resonance full width (in "normalized" frequency η) remains finite and equal to $2\mathscr{s}/a$. Referring to Fig. 4, we thus let

$$\xi = 2\ell + g, \quad \text{and} \quad \eta = j + h, \tag{55}$$

where ℓ and j are integers, and restrict $|g|$ and $|h|$ to both be less than or equal to $\mathscr{s}/a$. The factor of two is introduced since a scan of ξ for fixed η encounters raster elements half as frequently as does a scan of a line parallel to the ξ axis in the η-direction. We denote the power spectral density of the j^{th} resonance as a function of the displacement h from the resonance center j by $F_j(h) = F(j+h)$. Using (55), simplifying further to a slit with unit real transmission at the period's center, and examining the form of the ξ-integration (for both odd and even number of slits N), it may be written as

$$F_j(h) = \sum_{\ell=\frac{|j|+1-N}{2}}^{\frac{N-1-|j|}{2}} \int_{|h|-\frac{\mathscr{s}}{a}}^{\frac{\mathscr{s}}{a}-|h|} dg \frac{\exp\left[-\pi i\frac{\mu}{n}(j+h)(2l+g)\right]}{2} = \left[\sum_{\ell=\frac{|j|+1-N}{2}}^{\frac{N-1-|j|}{2}} \exp\left(-2\pi i\frac{\mu}{n}(j+h)l\right)\right] \times \left[\frac{1}{2}\int_{|h|-\frac{\mathscr{s}}{a}}^{\frac{\mathscr{s}}{a}-|h|} dg \exp\left(-\pi i\frac{\mu}{n}(j+h)g\right)\right]. \tag{56}$$

The ξ-integration has thus factored into a decoupled sum and integral, each of which may be evaluated independently. The integral is elementary and the sum is that of a geometric series. Thus, we may write the Fourier coefficient for the j^{th} harmonic (for $|h| \le \mathscr{s}/a$) as

$$F_j(h) = \left(\frac{\mathscr{s}}{a}-|h|\right)\operatorname{sinc}\left[\pi\frac{\mu}{n}(j+h)\left(\frac{\mathscr{s}}{a}-|h|\right)\right] \times \frac{\sin\left[\pi(N-|j|)\frac{\mu}{n}(j+h)\right]}{\sin\left[\pi\frac{\mu}{n}(j+h)\right]}, \tag{57}$$

where the function sinc(x) has its usual definition

$$\mathrm{sinc}(x) \equiv \frac{\sin(x)}{x}. \tag{58}$$

The imaginary part of the integrand of (53) is anti-symmetric under the interchange $\eta \leftrightarrow -\eta$ (or equivalently, $h \leftrightarrow -h$ and $j \leftrightarrow -j$) and cancels in the integration. Thus, with the substitutions of (55) and (18a), the inverse Fourier integral of the power spectrum becomes

$$|\tilde{\psi}(\tilde{x}_2)|^2 = \frac{2\mu}{n(R_1+R_2)} \sum_{j=0}^{N-1} \int_{-\frac{s}{a}\mathrm{sgn}(j)}^{\frac{s}{a}} dh \times \cos\left(2\pi\frac{\mu}{n}(j+h)\frac{\tilde{x}_2}{Ma}\right) F_j(h), \tag{59}$$

where the function sgn(x) has its usual definition

$$\mathrm{sgn}(x) \equiv \begin{cases} 1 & x>0 \\ 0 & x=0 \\ -1 & x<0. \end{cases} \tag{60}$$

For finite N our solution thus provides a finite number of finite width harmonics and a frequency shift, in sharp contrast to the infinite extent δ-function spectrum displayed by Cowley and Moodie's solution. Limiting the number of slits correspondingly limits the number of resonances, and thereby limits the spectral extent. The multiplicity m is similarly limited by N for self imaging to occur.

2.1 $N \to \infty$ Limit, Resonance Widths, and Frequency Shift

In Appendix B, we review Cowley and Moodie's result and find an image that is periodic with a period Ma, while in Sect. 1.2 we noted a basic image period shift to Ma/μ. We further noted that when $\mu = m$ is an integer, then the period Ma/m is commensurate with the period Ma and is compatible with the Fourier series of (B6). However, when μ is not an integer, then the periods Ma/μ and Ma are incommensurate, and the "broadened" Fourier series of (59) has an inherent period difference from the "sharp" Fourier series of (B6). How are we to resolve the apparent discrepancy?

In the $N \to \infty$ limit the solutions are indeed compatible. In this limit, we note that the Fraunhofer interference function (ratio of sines) in $F_j(h)$ approaches a comb function (infinite string of δ-functions) and the inverse Fourier transform becomes a convolution over $h+j$. The δ-functions occur for integer values $\mathscr{j} = (j+h)\mu/n$, or equivalently, at values $h = \mathscr{j}\frac{n}{\mu} - j$. The sum then becomes an infinite sum over $\mathscr{j}$, with the spatial frequencies of (59) being the same as those of (B6). Terms in the summation for which there exist no integers j and $\mathscr{j}$ such that the requirement $|h| \le \mathscr{s}/a$ is satisfied will vanish. Making these substitutions and performing the integration over h, the intensity may then be written as

$$|\tilde{\psi}(\tilde{x}_2)|^2 = \frac{2\mu}{n(R_1+R_2)} \sum_{\mathscr{j}=0}^{\infty} \left(\frac{\mathscr{s}}{a} - \left|\mathscr{j}\frac{n}{\mu} - j\right|\right) \times \cos\left(2\pi\mathscr{j}\frac{\tilde{x}_2}{Ma}\right) \mathrm{sinc}\left[\pi\mathscr{j}\left(\frac{\mathscr{s}}{a} - \left|\mathscr{j}\frac{n}{\mu} - j\right|\right)\right], \tag{61}$$

where j is now whatever integer makes the absolute value less than or equal to $\mathscr{s}/a$, and if none exists, then the associated term in the series is to be taken as zero.

Although not obvious from the forms of either (B6) or (61), the finite resonance width evident in (57) and (59) leaves traces latent in these results. A variation of λ (or equivalently, a variation of μ) will make the δ-functions occur at different values of h, which, in turn, will yield a variation of the Fourier coefficients of the "sharp" Fourier series. To see this, consider the multiplicity-m alias, whose dominant Fourier coefficient will be at $\mathscr{j} = m$. The associated δ-function will have a value of h given by the solution to the equations

$$\frac{\mu}{n}(n+h) = \left(\frac{m}{n} + \varepsilon\right)(n+h) = m. \tag{62}$$

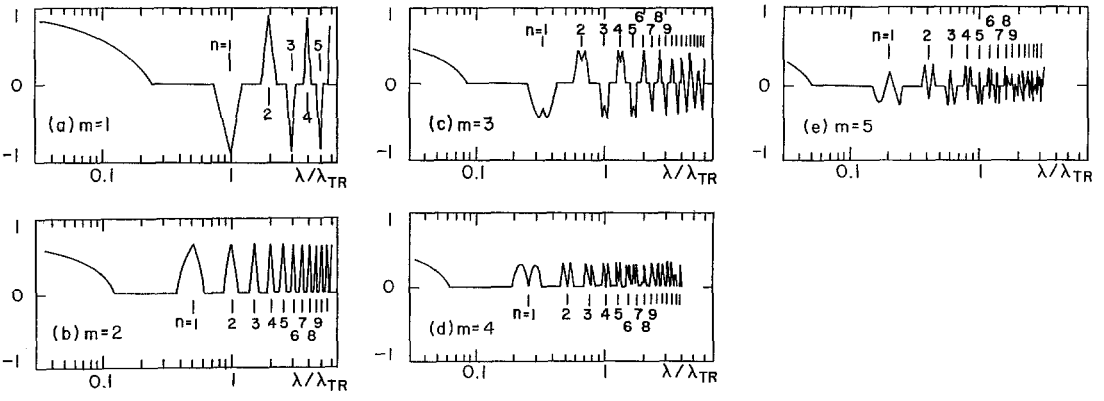

Fig. 5a–e. The dependence of the various spatial power spectrum Fourier coefficients on λ/λ_{TR} at $N = \infty$. Geometry ($R_1 = R_2 = 0.5$, $a = 10^{-5}$, $\mathscr{s}/a = 1/4$, $\lambda_{TR} = 4 \times 10^{-10}$, unit transmission slits) is held fixed while wavelength is varied, thereby displaying the finite-width resonance phenomena. **a** through **e** display, respectively, harmonic #1 (fundamental) showing $m = 1$ resonances through harmonic #5 showing $m = 5$ resonances. Tics mark associated n value at $\lambda/\lambda_{TR} = n/m$. Note the negative values for resonances with mn odd. Resonance widths are given by (65). **a** ($m = 1$) resonance peaks correspond to Talbot-von Lau fringes. Cusps on **d** ($n = 4$) resonances are due to the flat image formed by a sum of 4 self-images of $\mathscr{s}/a = 1/4$ slits. **e** ($m = 5$) resonances are due to piled up images for $\mathscr{s}/a > 1/m$. The resonances at $\lambda \to 0$ correspond to the geometric shadow pattern formed at $n = 0$

Furthermore, the m^{th} Fourier coefficient will vanish by the requirement $|h| \leq \lambda/a$, or equivalently at integral values of m and n for which the inequality

$$|\varepsilon| \leq \frac{m}{n^2}\frac{\lambda}{a} \tag{63}$$

does not hold. Hence, in the $N \to \infty$ limit the "sharp" Fourier coefficients of the intensity will display resonances for integral m and n for values of λ satisfying

$$\frac{\lambda_{\text{TR}}}{\lambda} = \frac{m}{n} + \varepsilon\,, \tag{64}$$

each with a full width as per [22]

$$\frac{\Delta\lambda}{\lambda} = \frac{2}{n}\frac{\lambda}{a}\,. \tag{65}$$

We demonstrate these assertions by a straight forward calculation. We have numerically evaluated (B6) for an infinite grating composed of rectangular slits, calculated the intensity for a single period, and then calculated its Fourier expansion. Various Fourier coefficients are plotted as a function of wavelength in Fig. 5. A string of coherent resonances is evident, each with a finite range of $\Delta\lambda$ given by (65). Additionally, we have repeated the procedure while varying the number of Fourier coefficients C_j in (B6) from 20 to 80. The resonance shapes and widths are essentially independent of the number of Fourier coefficients chosen in describing the amplitude.

3 Three Grating Lens-Free Interferometers

As an application of our results, we describe lens-free grating interferometers that employ a diffuse (spatially incoherent) monochromatic source and a diffuse detector. These are lens-free generalizations of a Lohmann and Silva interferometer [10] that works in a von Lau geometry. Given the foregoing discussion, the principles of operation for such an interferometer are quite straightforward to understand.

Three gratings with small λ/a and respective periods a_1, a_2, and a_3 are aligned with their slits parallel sequentially along an optic axis, with their planes spaced R_1 and R_2 apart. For spatially incoherent illumination of the first grating, consider the action of the middle (second) grating via our formulae with $a = a_2$. For illumination with narrowband light at a wavelength λ such that $\mu \approx m$ holds, then each point within a slit of the first grating acts as a source, incoherent from its neighboring source points. It produces a self-image of the second grating at the plane of the third grating, via the formulae given above. The third grating is selected to have a slit period a_3 equal a rational multiple of Ma_2/m, so that the transmission of the third grating forms a Moiré pattern with the image, i.e., the third grating has open slits at the positions of the self-images, and closed slits elsewhere. Different source points on the first grating will produce images on the third grating whose intensities will add incoherently. The first grating is selected so that it has a slit period that is a rational multiple of $Ma_2/(mM - m)$. With these choices for a_1 and a_3 all incoherently illuminated source points will image to form substantially similar patterns on the third grating and thereby maintain the visibility of the fringes on the third grating. Translation of any grating will then produce alternating bright and dark transmission. Any interspersed refractive elements, change in wavelength etc., will likewise cause a variation in the net transmission.

Earlier we discussed an example of such an interferometer in which the grating planes are equally spaced ($M = 2$) and the first and last gratings both have twice the period of the middle grating [16]. Many other configurations are possible. In the next section we give experimental results for another configuration.

Using the results of Sect. 2, the fringe visibility as a function of wavelength for fixed geometry, the visibility as a function of longitudinal grating displacement for fixed wavelength, or any other desired spectral property is readily calculated for such an interferometer. The effect of the first and last gratings on the intensity spectrum may be calculated by an appropriate filtering of the single source point image pattern.

3.1 Observation of Aliasing with Spatially Incoherent Light in a Lens-Free Three Grating Interferometer

In this section we describe the operation of one such interferometer discussed in the previous section. We thus report observation of aliasing using a lens-free interferometer employing both an uncollimated (and thus, spatially incoherent) monochromatic source and an uncollimated detector. Our experiment is readily reproducible in any undergraduate optics laboratory. Its geometry is shown in Fig. 6. Three sets of transmission gratings are mounted sequentially on an optical bench with their slits vertical and with their planes crudely parallel. The third is mounted on an optical translator to allow its horizontal displacement perpendicular to the bench axis, with its lateral position monitored by a machinist's dial gauge. The grating slits are aligned approximately parallel by shining a defocused, attenuated HeNe laser through the sequence, sighting into the beam over, and/or through, the various gratings, and then rotating the gratings about the optical axis so that the horizontal stripes produced by the grating's diffraction of the laser beam are visually parallel. A sodium vapor discharge lamp (borrowed from a UC-Berkeley undergraduate physics lab) is then placed close to the first grating. To minimize the effects of residual rotational misalignment of the gratings about the optical axis, the lamp's vertical extent is limited to 1.5 mm by masking the lamp with a horizontally oriented aluminum foil slit. The horizontal extent of the gratings and lamp is then about 25 mm.

An observation consists of sighting the lamp (from a convenient distance) through the three grating sequence and visually monitoring its intensity as the third grating is laterally

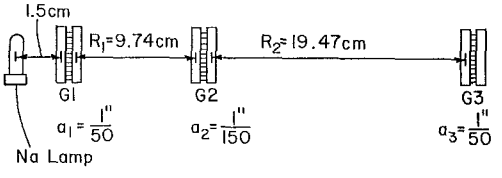

Fig. 6. Geometry of the experiment to demonstrate aliasing with spatially incoherent light in a lens-free three grating interferometer

translated. When the spacings between the grating planes are within about a millimeter of the positions predicted by our formulae (as measured by a meter stick), the visual appearance of the lamp's intensity will alternate between bright at the fringe maxima, and nearly extinguished at the minima. Given the logarithmic response of the human eye, the fringe contrast is thus very high. The dial gauge readings at these maxima and minima are then recorded.

Each grating consisted of a pair of identical inexpensive Ronchi "rulings" (purchased from Edmund's Scientific). Each "ruling" consists of a glass slide with a set of parallel black stripes reproduced on one face to produce a transmission grating with $\mathscr{s}/a \approx 0.5$. To produce a grating with $\mathscr{s}/a < 0.5$, the pair was held together with a rubber band with the emulsion sides in direct contact. We then rotationally adjusted the contacted gratings relative to each other to yield a uniform gray appearance, and then translationally adjusted their relative position to give the desired value for $\mathscr{s}/a$. The resulting $\mathscr{s}/a$ of the assembled grating was measured by measuring the net transmission of light through the assembly. A second method was also employed for measuring $\mathscr{s}/a$ and yielded the same value. It consisted of shining a HeNe laser through the grating and inferring $\mathscr{s}/a$ from the measured positions of the single slit diffraction intensity nulls.

The first and third grating had 50 stripes per inch ($a_1 = 0.508\,\text{mm}$), while the second had 150 ($a_2 = 0.169\,\text{mm}$). The three gratings were adjusted to $\mathscr{s}_1/a_1 = 0.006$, $\mathscr{s}_2/a_2 = 0.02$, and $\mathscr{s}_3/a_3 = 0.1$, respectively. To obtain maximum fringe contrast, the thickness b of each glass slide had to be compensated by decreasing the axial grating spacings by a corresponding small amount equal to b/n, where n is the index of refraction of the glass (assumed ≈ 1.5). After compensation, the effective spacing between the first and second gratings was $R_1 = 97.4\,\text{mm}$, and between the second and third was $R_2 = 194.7\,\text{mm}$, yielding $\varrho = 64.9\,\text{mm}$. These values were chosen to yield

$$\varrho = \frac{R_1 R_2}{R_1 + R_2} = \frac{4}{3}\frac{a_2^2}{\lambda_{\text{Na}}}, \qquad M = \frac{R_1 + R_2}{R_1} = 3,$$

$$\frac{R_2}{R_1} = 2, \qquad \frac{a_1}{a_2} = 3.$$

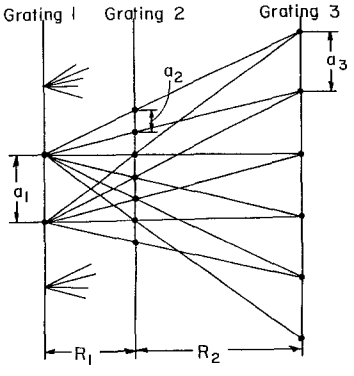

Fig. 7. The positions of the geometric shadows for our experimental arrangement. At our chosen wavelength and grating spacings there are *three* distinct images formed per geometric shadow period. A different choice of wavelength (and/or of $R_2 = 2R_1$) will form a different number of images per shadow period

The positions of the geometric shadows for this arrangement are shown in Fig. 7 and have a spacing equal to $a_3 = a_1$, which is that of a simple shadow Moiré, as well as that of the ($m = 1$) Talbot and von Lau effects. On the other hand, our formulae predict an alias multiplicity $m = 3$ and thus an image period of $Ma_2/m = a_2 = a_3/3$, i.e., one third of the geometric shadow spacing. By choosing a third grating period a_3 equal to a_1, upon lateral translation of this grating the observed brightness will then vary with the same period as the actual image period, and these two cases are easily distinguishable. Our dial gauge readings of the positions of maximal apparent brightness of nine successive fringes yield an observed fringe spacing $\approx a_1/3$, i.e., that of the aliased image and *not* that of the geometric shadow Moiré or the ($m = 1$) Talbot and von Lau effects.

4 Further Generalizations and Applications to Matter-Wave Interferometry

The analysis provided above may be extended to reflection gratings in a straightforward manner. It may also be generalized to allow the calculation of images produced by two-dimensional gratings as the analyses by Cowley and Moodie, and Winthrop and Worthington have done at $\mu = m = 1$ and $N = \infty$. Conceivably, it may also be extended to cylindrically symmetric gratings through the use of Bessel, Lommel, and Struve functions [2, Sect. 8.8].

An important application for lens-free three (and four) grating interferometers is to their use with matter-waves sand/or X-rays. Indeed, their use in atom interferometry provided our primary motivation for our pursuit of this work [11]. Both matter-waves associated with low energy neutral atoms and X-rays have wavelengths (0.01–10 nm) ideally suited for such devices using available micro-fabrication technology (micron to sub-micron range). If MKS units are used for the examples of Figs. 3 and 5, the values are typical for a neutral atom matter-wave interferometer.

Low energy atoms (and soft x-rays) have negligible transmission range through matter. Thus, the use of vacuum apertures provides an attractive option for the construction of aberration-free optics for these waves. Separated beam atom interferometry [12] current practice has a low throughput flux dominantly limited by its collimation requirements. Use of generalized Talbot-von Lau fringes and the interferometer geometries discussed herein remove these limits and promise very high throughput fluxes with arbitrarily wide collimators. Such interferometers thus have a wide variety of applications for both X-rays and matter waves, including (with gratings periodic in 2D and/or 1D) a phase sensitive imaging device and/or a narrow-band interference filter. The grating dimensions required for the interferometers described herein are far easier (by more than an order of magnitude) to accommodate than those required for separated-beam atom interferometer configurations. Alignment of low velocity (long wavelength) interferometers is further facilitated by the presence of the high velocity ($\mu \to 0$), geometric shadow Moiré. Conceivably, given the slow dependence of the required grating period upon the wavelength, a $\propto \lambda^{1/2}$, even hard x-ray interferometry may be performed via the above techniques with gratings formed from suitably selected crystals (e.g., with a

homogeneous lattice, or with one composed of alternating high and low 2 species).

For matter-wave interferometry, it is straightforward to show that long range potentials (such those due to gravity and rotation), when acting on the matter-waves associated with massive particles within an interferometer, simply shift the diffraction pattern by a classical displacement. Thus, our interferometers can be useful probes for measuring effects associated with such potentials. When the open topology of the propagation is also considered, quantum mechanical topological phase shifts will modify these effects and must also be accounted for. A consideration of these modifications will be the subject of future work [23].

Appendix A. The Talbot and von Lau Effects

In 1836, Talbot made a remarkable observation [13]. A diagram of his apparatus is shown in Fig. 8a. In this experiment a beam of monochromatic light is focused parallel with a lens and directed through a coarse periodic binary transmission grating (Ronchi ruling). A second identical grating is then placed in the transmitted light and aligned so that its slit plane and slits are parallel to those of the first grating. An observation consisted of varying the separation between the two planes and the lateral position of the second grating (in a direction perpendicular to the grating slits) while monitoring the light transmitted by the pair.

When the spacing between the gratings is zero (i.e., when they are in contact with each other), a simple Moiré variation of the transmitted intensity (triangular in shape and readily explained) is produced by lateral displacements of the second grating. When the gratings are slightly separated, the Moiré transmission blurs. However, at grating spacings equal to integral multiples of a characteristic length the Moiré fringes reappear at full visibility! Moreover, at odd multiples of this length, the Moiré fringe phase is displaced by 180°. This effect remained unexplained until 1881, whereupon Lord Rayleigh identified the characteristic length to be given by $2a^2/\lambda$, where a is the grating period and λ is the wavelength of the light [14].

In 1936, von Lau performed a different experiment [15]. A diagram of his apparatus is shown in Fig. 8b. In von Lau's experiment an extended (spatially incoherent) monochromatic source illuminates a separated pair of identical gratings, one through the next. Light exiting the second grating then passes through a lens. At the focal distance of the lens a screen is used to observe thus magnified patterns thus produced. Again, when the gratings are spaced at integral multiples of a characteristic length, exact images of the gratings are formed on the screen.

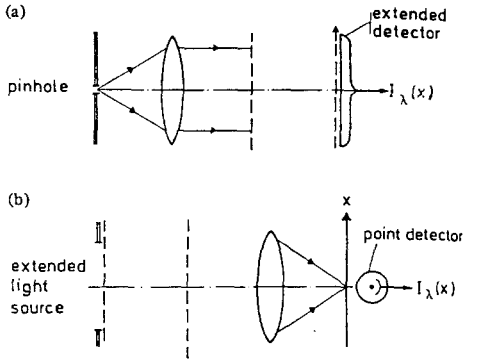

Fig. 8. a Diagram of Talbot's apparatus. b Diagram of von Lau's apparatus. (Figures after Jahns and Lohmann)

Appendix B. Treatment by Cowley and Moodie

The treatment by Cowley and Moodie for $N = \infty$ is readily generalized to include complex transmission functions. In this Appendix, we review their argument using our notation. We simplify their treatment of two-dimensional gratings to our domain of one dimension. We start from (5), which we may write in terms of λ_{TR} via (14) as

$$\tilde{\psi}(\tilde{x}_2) = \frac{a e^{i\phi_0}}{(\lambda R_1 R_2)^{1/2}} \int_{-\infty}^{\infty} du\, t(au) \times \exp\left[i\pi \frac{\lambda_{TR}}{\lambda}\left(u - \frac{\tilde{x}_2}{aM}\right)^2\right], \tag{B1}$$

and where we have made the variable change $u = x_1/a$. We specify the grating to be infinite in extent and periodic with period a, whereupon we express its amplitude transmission function by the Fourier series

$$t(x_1) = \sum_{j=-\infty}^{\infty} C_j \exp\left(\frac{i2\pi j x_1}{a}\right). \tag{B2}$$

Using (B2), we may write (B1) as

$$\tilde{\psi}(\tilde{x}_2) = \frac{a e^{i\phi_0}}{(\lambda R_1 R_2)^{1/2}} \sum_{j=-\infty}^{\infty} C_j \int_{-\infty}^{\infty} du \times \exp\left\{i\pi\left[2ju + \frac{\lambda_{TR}}{\lambda}\left(u - \frac{\tilde{x}_2}{aM}\right)^2\right]\right\}. \tag{B3}$$

By completing the square of the argument of the exponential factor in (B3), the amplitude may be written as

$$\tilde{\psi}(\tilde{x}_2) = \frac{a e^{i\phi_0}}{(\lambda R_1 R_2)^{1/2}} \times \sum_{j=-\infty}^{\infty} C_j \exp\left\{-i\pi\left[\frac{\lambda}{\lambda_{TR}} j^2 - 2j\frac{\tilde{x}_2}{aM}\right]\right\} \times \int_{-\infty}^{\infty} du \exp\left[i\pi \frac{\lambda_{TR}}{\lambda}\left(u + \frac{\lambda}{\lambda_{TR}} j - \frac{\tilde{x}_2}{aM}\right)^2\right] \tag{B4}$$

wherein we have moved the factor that contains no u dependence outside of the integral. Upon making the variable change

$$u' = u + \frac{\lambda}{\lambda_{TR}} j - \frac{\tilde{x}_2}{aM}, \tag{B5}$$

the integration over u' is recognized as a Fresnel integral and is immediately performed, yielding

$$\tilde{\psi}(\tilde{x}_2) = \frac{e^{i(\phi_0+\pi/4)}}{(R_1+R_2)^{1/2}} \times \sum_{j=-\infty}^{\infty} C_j \exp\left(-i\pi \frac{n}{\mu} j^2\right) \times \exp\left(i2\pi j \frac{\tilde{x}_2}{aM}\right), \quad \text{(B6)}$$

where (as per (18) above) we use the definition $\lambda/\lambda_{TR} \equiv n/\mu$, and where the function sgn(j) has its usual definition as per (58). Since (B6) has the form of a Fourier series, the amplitude is a periodic function of $\tilde{x}_2$, with period Ma.

Cowley and Moodie found special significance only in cases with $\mu = 1$. Then, when λ/λ_{TR} is an even integer n, the first exponential factor is one, and the amplitude becomes a magnified exact image of the grating via (B2). When n is an odd integer, Cowley and Moodie recognized the resulting intensity to represent a magnified exact image of the grating shifted by one half period. What Cowley and Moodie missed in discussing this formula is that when μ is any integer m, then the amplitude (except for an extraneous quadratic phase factor) is a periodic function of $\tilde{x}_2$, with period Ma/m. The existence of m aliased intermediate images for $\mu = m \neq 1$ was recognized by Gori [8] and by Sudol and Thompson [9], and was experimentally demonstrated by Gori for a von Lau configuration. None of these authors commented that the aliased image *amplitudes* would be exact self-image amplitudes of an individual grating single period complex amplitude transmission. Gori did, however, suggest that self-imaging will occur for slits with non-zero s/a for a finite but unspecified range of $\mu \neq m$. The analysis presented above proves these assertions to be true and further specifies the range in detail [24].

Appendix C. Generalized Quadratic Fresnel Zone Plates

Jahns and Lohman [4] have pointed out that the Talbot and von Lau effects are reciprocals of each other in which the roles of source and detector are interchanged. In the von Lau geometry (i.e., for $R_2 = \infty$), they give a simple geometric construction that allows these effects to be understood as the action of a generalized quadratic Fresnel zone plate. They consider the path lengths for light passing from a source point on the axis through each slit to a detection point on the axis. They note that at the Talbot-Rayleigh condition ($m = 1$) the path length through each slit is then equal to the square of an integer times the wavelength. Since the square of an integer is again an integer, focusing to a spot on the axis in a fashion similar to that of a Fresnel zone plate is expected. Their simple picture is readily generalized to explain a full array of aliased foci for $n = 1$, for any integer m (as defined by (18)), and for finite R_1 and R_2. We use the fact that the quantity $j(mj \pm \ell)/2$ is also an integer when j is any integer, whenever the requirement holds that ℓ is any odd integer when m is odd, and ℓ is any even integer when m is even. A focusing/aliasing geometry results (in our notation) whenever $\tilde{x}_2/M = (x_2 - x_{20})/M = \ell/(2m)$ holds. When N is odd, the j^{th} slit is then spaced from the axis as $x_{1j} = ja$, where j is an integer. The difference in path lengths through slits $j = f$ and $j = g$ to a spot at x_2 is then given by

$$\Delta_f - \Delta_g = \frac{x_{1f}^2 - x_{1g}^2}{2\varrho} = \frac{a^2}{2\varrho}\left[f\left(f - \frac{\ell}{m}\right) - g\left(g - \frac{\ell}{m}\right)\right]. \quad \text{(C1)}$$

Rewriting (18a) for $\mu = m$ and $n = 1$ as

$$\frac{a^2}{\varrho} = m\lambda, \quad \text{(C2)}$$

the path length difference becomes

$$\Delta_f - \Delta_g = \lambda\left[\frac{f(mf - \ell) - g(mg - \ell)}{2}\right]. \quad \text{(C3)}$$

Thus, for the above restrictions on ℓ relative to m, the path length difference is always a multiple of the wavelength. The array of foci and associated half-period shift are thus explained. The case where N is even follows similarly, since the slits are then spaced from the axis as $y_j = (j + 1/2)a$, whereupon the the above result obtains as before, but with ℓ decremented by one, i.e., with the iopposite requirement on ℓ such that ℓ is now any odd integer when m is even and ℓ is any even integer when m is odd. The situation for $n \neq 1$ is more complicated and evidently does not lend itself to this simple visualization in terms of a quadratic zone plate.

References

1. See, for example, J.W. Goodman: *Introduction to Fourier Optics* (McGraw-Hill, New York 1968) Chaps. 3, 4
2. M. Born, E. Wolf: *Principles of Optics* (Pergamon, Oxford 1987) Chap. 8
3. See, for example, G.L. Rogers: Br. J. Appl. Phys. **14**, 657 (1963)
   A. Zeilinger et al.: Revs. Mod. Phys. **60**, 1067 (1988)
   Q.A. Turchette, D.E. Pritchard, D.W. Keith: Numerical Modeling of a Multiple Grating Interferometer. MIT preprint (1991)
4. J. Jahns, A.W. Lohmann: Opt. Commun. **28**, 263 (1979)
5. See also J.T. Winthrop, C.R. Worthington: J. Opt. Soc. Am. **55**, 373 (1965); **56**, 588 (1966)
   E.A. Hiedemann, Breazeale: J. Opt. Soc. Am. **49**, 372 (1959)
   B.J. Chang, R. Alferness, E.N. Leith: Appl. Opt. **14**, 1592 (1975)
6. J.M. Cowley, A.F. Moodie: Proc. Phys. Soc. B **70**, 486, 497, 505 (1970)
7. Cowley and Moodie comment: "In fact it is observed that with gratings of this type a large number of sharp and frequently complicated out-of-focus patterns are generated." They called these intermediate images Fresnel images. In the present work we identify these images as those for which m is an integer not equal to one. They reserved the name Fourier images as those for which $m = 1$ holds. Although the existence of $m \neq 1$ images is indeed manifest in their formulae, the regularity of the patterns was not recognized by them, nor obvious from their formulae. Thus, they comment (p. 499) that "No obvious relationship exists between the positions of the delta functions and the maxima and minima of the real and imaginary parts of the Fourier transform of the Fresnel wave function."...

8. F. Gori: Opt. Commun. **31**, 4 (1979)
9. R. Sudol, B.J. Thompson: Opt. Comm. **31**, 105 (1979)
10. A.W. Lohman, D.E. Silva: Opt. Commun. **2**, 413 (1971)
11. J.F. Clauser: Rotation, Acceleration, and Gravity Sensors Using Quantum Mechanical Matter-Wave Interferometry with Neutral Atoms and Molecules. United States Patents #4, 874, 942 (1987, 1989) and #4, 992, 656 (1989, 1991); Physica B **151**, 262 (1988); Ultra-Sensitive Inertial Sensors via Neutral-Atom Interferometry, in *Relativistic Gravitational Experiments in Space*. NASA Conf. Publ. 3046, ed. by R.W. Hellings (1989)
12. See B.G. Levi: Physics Today **44**, #7, 17 (1991) for a discussion of the current experimental status
13. H. Talbot: Philos. Mag. **9**, 401 (1836)
14. Lord Rayleigh: Philos. Mag. **11**, 196 (1881)
15. E. Lau: Ann. Phys. **6**, 417 (1948)
16. M.W. Reinsch, J.F. Clauser: Bul. Amer. Phys. Soc. **36**, 1312 (1991)

Notes added in proof:

17. The term *image* is used loosely here. Although a diffraction grating is not *per se* an imaging device, nonetheless sharp self-images are indeed formed
18. See, however, the excellent review: K. Patorski, *Progress in Optics XXVII*, ed. by E. Wolf (Elsevier, Amsterdam 1989) pp. 1–108 and the extensive reference list therein
19. Whereas we are considering a one-dimensional grating and cylindrical wave propagation, (5) is obtained by integrating the usual scalar diffraction theory formula [1,2] over the y-dimension
20. Some authors have a give a more complicated expression for the aliasing multiplicity, because they use integers that are defined differently. For example [17, p. 26]. Patorski's expression $2\nu + 2N/n$ equals our n/m
21. Note, however, that E. Menzel and Ch. Menzel, Optik **4**, 22 (1948), suggest that a minimum of 17 slits are required for such effects to be observed
22. Note that A.P. Smirnov, Opt. Spectrosc. **46**, 319 (1979), earlier gave an order-of-magnitude estimate for this effect
23. Although this work considers only scalar waves, vector and/or spinor diffraction theory is required for a discussion of topological phase shifts associated with spin and/or polarization
24. Since we became aware of [17] only after completion of this manuscript, it appears that some of these features were (at least qualitatively) known to the authors of references cited therein. We thus apologize to these authors for any omitted citations