JOURNAL OF MODERN OPTICS, 1996, VOL. 43, NO. 10, 2139–2164

Integer, fractional and fractal Talbot effects

M. V. BERRY and S. KLEIN
H. H. Wills Physics Laboratory, Tyndall Avenue,
Bristol BS8 1TL, England

(*Received 21 February 1996; revision received 12 March 1996*)

Abstract. Self-images of a grating with period a, illuminated by light of wavelength λ, are produced at distances z that are rational multiples p/q of the Talbot distance $z_T = a^2/\lambda$; each unit cell of a Talbot image consists of q superposed images of the grating. The phases of these individual images depend on the Gauss sums studied in number theory and are given explicitly in closed form; this simplifies calculations of the Talbot images. In `transverse' planes, perpendicular to the incident light, and with $\zeta = z/z_T$ irrational, the intensity in the Talbot images is a fractal whose graph has dimension $\frac{3}{2}$. In `longitudinal' planes, parallel to the incident light, and almost all oblique planes, the intensity is a fractal whose graph has dimension $\frac{7}{4}$. In certain special diagonal planes, the fractal dimension is $\frac{5}{4}$. Talbot images are sharp only in the paraxial approximation $\lambda/a \to 0$ and when the number N of illuminated slits tends to infinity. The universal form of the post-paraxial smoothing of the edge of the slit images is determined. An exact calculation gives the spatially averaged non-paraxial blurring within Talbot planes and defocusing between Talbot planes. Similar calculations are given for the blurring and defocusing produced by finite N. Experiments with a Ronchi grating confirm the existence of the longitudinal fractal, and the transverse Talbot fractal at the golden distance $\zeta = (3 - 5^{1/2})/2$, within the expected resolutions.

1. Introduction

What is the field beyond a coherently illuminated diffraction grating? This simplest of questions in wave optics still holds some surprises, whose origin lies in Talbot's [1] remarkable discovery, in 1836, of a series of self-images in planes beyond the grating. In recent years his effect has received much attention, both as a fundamental optical phenomenon and because of its applications [2]. Rayleigh [3] explained the Talbot images in planes whose distances z from the grating are even multiples of the *Talbot distance* $z_T = a^2/\lambda$, where a is the period of the grating and λ the wavelength of the light. Much later, it was observed [4] and explained with the aid of the paraxial (Fresnel) approximation [5] that images form at all rational multiples of z_T, namely $z = (p/q)z_T$, where p and q are coprime integers. Within each unit cell of the image plane, these fractional Talbot images consist of q equally spaced copies of the transmission function of the grating, which superpose coherently when they overlap.

Here we make four contributions to this subject. First (section 2), we show that the phases of the Talbot images possess an irreducible arithmetic complexity, arising from the Gauss sums studied in number theory, which can be evaluated explicitly in closed form. These phases govern the intensities of the high-order (i.e.

0950–0340/96 $12·00

large q) fractional images, because these overlap, even for narrow slits. Related to the connection with Gauss sums is the mathematically interesting fact that within the paraxial approximation the wave describing the essence of the Talbot effect is a Jacobi theta function on the natural boundary of its region of convergence. Therefore the intensity variation, for a grating whose slits have sharp edges, is highly singular, both in and between Talbot planes.

Second (section 3), we find that for gratings with sharp-edged slits (e.g. Ronchi gratings) the light intensity has an amazing fractal structure. By studying the limit of sequences of fractional Talbot distances $z_T p/q$ where p/q tends to an irrational number, we argue that for almost all irrational image planes the graph of the intensity is a continuous curve whose fractal dimension [6] is $\frac{3}{2}$. It follows that the nature of the images depends sensitively on z; therefore the Talbot effect can be regarded as an optical `arithmoscope', revealing the number-theoretic nature (e.g. rational or irrational) of $z/z_T$. However, these $\frac{3}{2}$-dimensional fractals are special. In typical non-transverse planes, for example the `longitudinal' planes parallel to the incident light and the slits of the grating, the graph of intensity has the less familiar dimension $\frac{7}{4}$. For particular diagonal planes, the graph has dimension $\frac{5}{4}$.

Each integer or fractional Talbot image arises from constructive interference between light from all slits in the grating, but only in the paraxial approximation does this give rise to perfectly sharp images. Our third contribution (section 4) is to give a post-paraxial theory for the imperfections of the Talbot images, complementing a recent numerical study [7]. It is interesting that, although the Talbot images are not formed by geometrical focusing, the condition for the emergence of the paraxial limit, in which they are sharp, is $\lambda/a \to 0$, resembling conditions for the emergence of the sharp caustic singularities of geometrical optics. Post-paraxial effects blur the sharp edges of the images and wipe out the finest scales of the fractals, according to a universal diffraction function with fringes whose lateral scale is $(z/z_T)^{1/4}(\lambda a)^{1/2}$. We also give an exact non-paraxial calculation of the spatially averaged sensitivity of the light intensity to variations in and between Talbot planes.

Even within the Fresnel approximation, the sharpness of the Talbot images requires a grating with infinitely many slits. Our fourth contribution (section 5) is to quantify the effect of a large but finite number N of slits. The (unsurprising) results is that the sharp edges of the Talbot images, and the fractals, are blurred by fringes with lateral scale $a(z/z_T)/N$. The Talbot effect for N narrow slits is related to a previous study [8] of incomplete Gauss sums, and the associated hierarchical generalizations of Cornu spirals (`curlicues'). We also calculate, within the paraxial approximation but with finite N, the spatially averaged sensitivity of the light intensity to variations in and between Talbot planes.

The main new physical effects predicted by this work are the fractal intensity distributions beyond the grating. In section 6 we describe experiments with a Ronchi grating (binary grating with slit width $a/2$), demonstrating the transverse Talbot fractal (at the golden ratio $((3-5^{1/2})/2)z_T$), and also the longitudinal Talbot fractal.

In this paper we consider only one-dimensional gratings, illuminated by monochromatic light. The two-dimensional case is considered in many of the references cited; if the ratio of the grating periods in the two directions is rational, the theory is essentially the same as in one dimension, although the details can get complicated. Illumination by white light (as in Talbot's [1] original experiment)

gives richly coloured images, especially close to the grating; we do not pursue this aspect here.

A class of different physical effects with the same underlying mathematics is the integer and fractional revivals of electron wavepackets in Rydberg atoms [9–11]. Instead of distance from the grating, there is time in the evolution of the electron wave; the periodic coordinate across the grating and in the Talbot observation planes is replaced by the angle round the ring near which the electron wave is confined; instead of the wave immediately beyond the grating, there is the initial electron wave packet. It is possible that the analysis presented here can be usefully applied to Rydberg revivals.

2. Gauss sums and phases of fractional Talbot images

A plane wave of light with wavenumber $k = 2\pi/\lambda$, travelling in the z direction, is transmitted by a grating that varies in the transverse direction x and then propagates freely in the space beyond. It is convenient to employ dimensionless coordinates ξ and ζ, scaled relative to x and z by the grating period and the Talbot distance, namely

$$\xi \equiv \frac{x}{a}, \qquad \zeta \equiv \frac{z}{z_{\mathrm{T}}} = \frac{z\lambda}{a^2}. \tag{1}$$

Let the transmission function in the unit cell $|x| < a/2$ be $g(\xi)$; for example, a Ronchi grating, that is a binary grating with open slit width $a/2$, has

$$g(\xi) = 1 \quad (|\xi| \leqslant \tfrac{1}{4}), \qquad g(\xi) = 0 \quad (|\xi| > \tfrac{1}{4}). \tag{2}$$

At the point (ξ, ζ) beyond the grating, the scalar wave (that is neglecting polarization effects, for which see [12]) is

$$\Psi(\xi, \zeta) = \int_{-1/2}^{1/2} \mathrm{d}\xi' g(\xi') \psi_{\mathrm{comb}}(\xi - \xi', \zeta), \tag{3}$$

where $\psi_{\mathrm{comb}}(\xi, \zeta)$ is the wave transmitted by a grating consisting of a comb of Dirac δ slits.

The comb wave is conveniently written as a superposition of diffracted beams. Exactly,

$$\psi_{\mathrm{comb}}(\xi, \zeta) = \sum_{n=-\infty}^{\infty} \exp(2\pi \mathrm{i} n \xi) \exp\left\{2\pi \mathrm{i} \zeta \left(\frac{a}{\lambda}\right)^2 \left[1 - \left(\frac{n\lambda}{a}\right)^2\right]^{1/2}\right\}, \tag{4}$$

where the square root is positive imaginary for the evanescent waves n for which its argument is negative. This formula applies to an infinite grating; the effect of finite numbers of slits will be studied in section 5. Later it will be convenient to write equations (3) and (4) in the form

$$\Psi(\xi, \zeta) = \sum_{n=-\infty}^{\infty} g_n \exp(2\pi \mathrm{i} n \xi) \exp\left\{2\pi \mathrm{i} \zeta \left(\frac{a}{\lambda}\right)^2 \left[1 - \left(\frac{n\lambda}{a}\right)^2\right]^{1/2}\right\}, \tag{5}$$

where g_n is the nth Fourier coefficient of the transmission function, that is

$$g_n \equiv \int_{-1/2}^{1/2} \mathrm{d}\xi g(\xi) \exp(-2\pi i n\xi). \tag{6}$$

For the Ronchi grating (2),

$$g_0 = \tfrac{1}{2}, \qquad g_{n=2k+1} = \frac{(-1)^k}{\pi(2k+1)}. \tag{7}$$

Now we make the central approximation that the propagation is paraxial and expand the square root to the term in n^2. This gives the comb wave in terms of the *paraxial propagator* $\psi_p(\xi, \zeta)$, defined by

$$\psi_{comb}(\xi, \zeta) \approx \exp(ikz)\, \psi_p(\xi, \zeta), \tag{8}$$

where

$$\psi_p(\xi, \zeta) = \sum_{n=-\infty}^{\infty} \exp(2\pi i \xi n - i\pi\zeta n^2).$$

The analogue of equation (5), for an arbitrary grating, is the paraxial scalar wave (i.e. the paraxial approximation to equation (5) with the factor $\exp(ikz)$ removed)

$$\Psi_p(\xi, \zeta) = \sum_{n=-\infty}^{\infty} g_n \exp(2\pi i \xi n - i\pi\zeta n^2). \tag{9}$$

The function ψ_p in equation (8) is a Jacobi theta function (section 16.27 of [13]) but the sum over n would converge only if ζ had a negative imaginary part. Real ζ, as in the Talbot effect, corresponds to the highly singular situation of a theta function on its natural boundary. Nevertheless, this gives sensible results in combination with the integral (3) indeed, the singularities are responsible for the sharp images that are the essence of the Talbot effect.

We note that the Poisson sum formula, which we shall employ extensively in what follows, reproduces the δ-comb initial condition

$$\psi_p(\xi, 0) = \psi_{comb}(\xi, 0) = \sum_{n=-\infty}^{\infty} \exp(2\pi i \xi n) = \sum_{m=-\infty}^{\infty} \delta(\xi - m). \tag{10}$$

In section 4 we shall give a detailed discussion of the validity of the paraxial approximation and show that its apparent contradiction with the assumption of an infinite grating disappears for small λ/a. Anticipating this result, we now turn to a systematic study of the paraxial propagator.

It is easy to confirm from equation (8) that $\psi_p(\xi, \zeta)$ possesses the following symmetries:

$$\begin{aligned} \psi_{\mathrm{p}}(\xi+1,\zeta) &= \psi_{\mathrm{p}}(-\xi,\zeta) = \psi_{\mathrm{p}}(\xi,\zeta), \\ \text{so}\quad \psi_{\mathrm{p}}(1-\xi,\zeta) &= \psi_{\mathrm{p}}(\xi,\zeta), \\ \psi_{\mathrm{p}}(\xi,\zeta+1) = \psi_{\mathrm{p}}(\xi+\tfrac{1}{2},\zeta), &\qquad \psi_{\mathrm{p}}(\xi,-\zeta) = \psi_{\mathrm{p}}^{*}(\xi,\zeta), \\ \text{so}\quad \left|\psi_{\mathrm{p}}(\xi,1-\zeta)\right|^2 &= \left|\psi_{\mathrm{p}}(\xi+\tfrac{1}{2},\zeta)\right|^2. \end{aligned} \tag{11}$$

Therefore it suffices to know the paraxial propagator for $|\zeta| \leqslant \frac{1}{2}$, $|\xi| \leqslant \frac{1}{2}$. For gratings with sufficient symmetry, for example the Ronchi grating (2), these symmetries are inherited by the (paraxial) intensity $|\Psi_{\mathrm{p}}|^2$ given by equation (9).

Fractional Talbot images occur in planes where z is a rational multiple of z_{T}, that is $\zeta = p/q$, where p and q are coprime integers. Then in equation (8) we split the sum into groups of q terms by defining

$$n = ql + s \quad (-\infty < l < \infty,\ 1 \leqslant s \leqslant q). \tag{12}$$

Thus

$$\psi_{\mathrm{p}}\left(\xi,\frac{p}{q}\right) = \sum_{s=1}^{q} \exp\left[\mathrm{i}\pi\left(2s\xi - \frac{p}{q}s^2\right)\right] \sum_{l=-\infty}^{\infty} \exp\left[\mathrm{i}\pi(2lq\xi - l^2qp - 2lsp)\right]. \tag{13}$$

In the second exponential the term $2lsp$ can be neglected because it contributes a factor unity. The term in l^2 can be simplified, because

$$\exp(-\mathrm{i}\pi l^2 qp) = (-1)^{l^2 qp} = (-1)^{lqp} = \exp(-\mathrm{i}\pi l q e_p), \tag{14}$$

where $e_p \equiv 0\,(1)$ if p is even (odd).

The observation (14) is the mathematical heart of the Talbot effect, because it renders the exponent in the second exponential in equation (13) linear in l, so that the Poisson formula (10) can be applied to give the paraxial propagator as a δ comb for any rational ζ and not just $\zeta = 0$. The result is

$$\psi_{\mathrm{p}}\left(\xi,\frac{p}{q}\right) = \frac{1}{q^{1/2}} \sum_{n=-\infty}^{\infty} A(n;q,p)\,\delta\left(\xi' - \frac{n}{q}\right), \tag{15}$$

where

$$\xi' \equiv \xi - \tfrac{1}{2}e_p \tag{16}$$

and

$$A(n;q,p) \equiv \frac{1}{q^{1/2}} \sum_{s=1}^{q} \exp\left(\mathrm{i}\frac{\pi}{q}\left[2s(n+\tfrac{1}{2}qe_p) - ps^2\right]\right). \tag{17}$$

It is easy to show (by an elementary change of variables in the double sum) that $|A(n;q,p)|^2 = 1$. Therefore the A are pure phase factors, that is

$$A(n;q,p) = \exp\left[\mathrm{i}\Phi(n;q,p)\right]. \tag{18}$$

From equation (3), we now obtain the paraxial wave beyond a general grating:

$$\Psi_{\mathrm{p}}\left(\xi\ \frac{p}{q}\right) = \frac{1}{q^{1/2}} \sum_{n=-\infty}^{\infty} g\left(\xi - \frac{n}{q}\right) \exp\left[\mathrm{i}\Phi(n;q,p)\right]. \tag{19}$$

This equation captures the physics of the Talbot effect; in each unit cell (e.g. $0 \leqslant \xi \leqslant 1$) of the plane $\zeta = p/q$, images of the grating slits are reproduced with spacing $1/q$ and intensity reduced by $1/q$; if p is even, the $n = 0$ image is at $\xi = 0$ and, if p is odd, the $n = 0$ image is laterally shifted [14–16] and lies at $\xi = \frac{1}{2}$.

Now we observe that the quantities $A(n;q,p)$ are the Gauss sums studied in classical number theory, which can be evaluated in closed form. In the commonest form of these sums [17], the term linear in s in the exponent of equation (17) is absent, but the same techniques can be applied when this term is included; an elementary derivation was given by Hannay in Appendix A of [18]. The result is that the phases can be represented in explicit form as

$$\Phi(n;q,p) = \begin{cases} \pi\left\{\frac{1}{4}(q-1) + \dfrac{p}{q}(p\backslash q)^2 n^2 + \dfrac{1}{2}\left[1 - \left(\dfrac{p}{q}\right)\right]\right\} & (p \text{ even}, q \text{ odd}), \\ \pi\left\{\frac{1}{4}p + \dfrac{p}{q}(p\backslash q)^2 (n+\frac{1}{2}q)^2 + \dfrac{1}{2}\left[1 - \left(\dfrac{q}{p}\right)\right]\right\} & (p \text{ odd}, q \text{ even}), \\ \pi\left\{\frac{1}{4}(q-1) + \dfrac{4p}{q}(4p\backslash q)^2 (n+\frac{1}{2}q)^2 + \dfrac{1}{2}\left[1 - \left(\dfrac{p}{q}\right)\right]\right\} & (p \text{ odd}, q \text{ odd}). \end{cases} \tag{20}$$

Here, $(p\backslash q)$ is the integer inverse of $p \bmod q$, given by

$$(p\backslash q) = p^{(\phi_{\mathrm{E}}(q)-1)} \bmod q, \tag{21}$$

where $\phi_{\mathrm{E}}(q)$ denotes Euler's totient function (number of positive integers less than q that are relatively prime to q), and $\left(\dfrac{p}{q}\right)$ is the Jacobi symbol, which takes the value +1 or − 1, and is defined as the product of Legendre symbols $\left(\dfrac{p}{s}\right)$ for all prime factors s of q, these in turn being defined as

$$\left(\frac{p}{s}\right) \equiv \begin{cases} +1 & \text{if there is an integer } m \text{ such that } m^2 = p \bmod s, \\ -1 & \text{if there is no integer } m \text{ such that } m^2 = p \bmod s. \end{cases} \tag{22}$$

These phases are important because their n dependence determines the intensity of the superposition of overlapping grating images. For the Ronchi grating (2), overlap occurs whenever $q > 2$.

With these formulae it is easy to compute the intensities of the Talbot images in any plane $\zeta = p/q$, for any grating function $g(\xi)$ (for example using the program *Mathematica* [19]), thereby extending previous computations (for example [20, 21]) for special cases. Figure 1 shows some examples.

If the incident wave is not plane but comes from a source at a distance R in front of the grating, then, as is well known (for example [22]), all the preceding formulae apply provided that ξ and ζ are replaced by ξ_R and ζ_R, where

$$\frac{\xi}{\xi_R} = \frac{\zeta}{\zeta_R} = 1 + \frac{z}{R}. \tag{23}$$

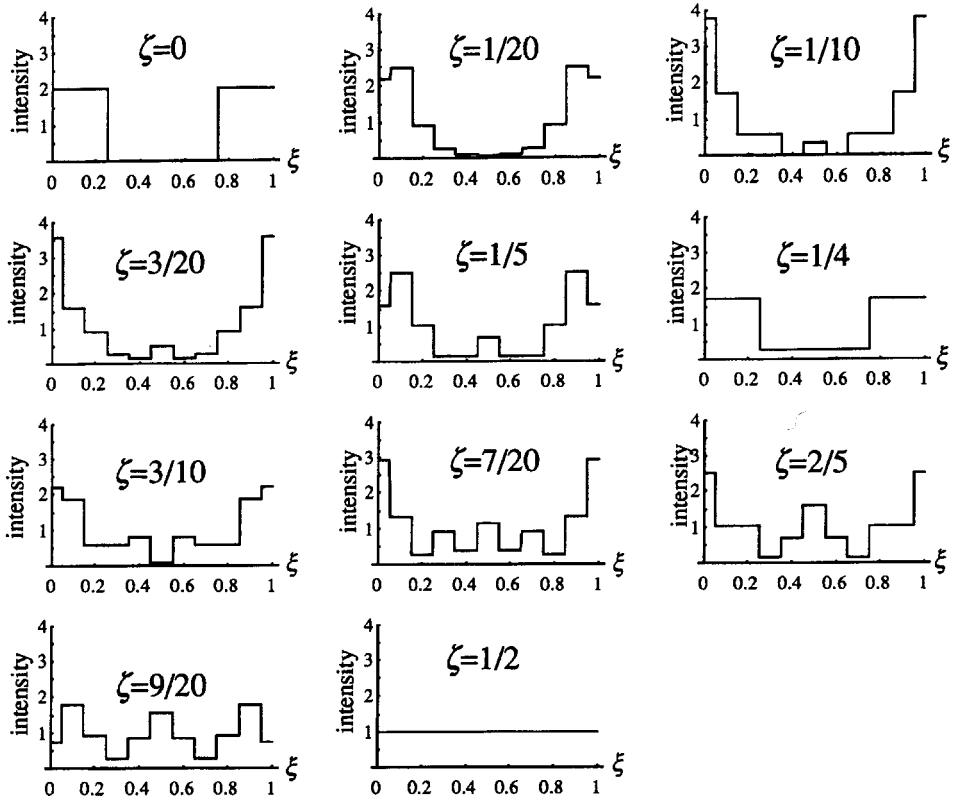


Figure 1. Intensities in one unit cell of the Talbot planes $\zeta = n/20$ ($0 \leqslant n \leqslant 10$), for the Ronchi grating (2).

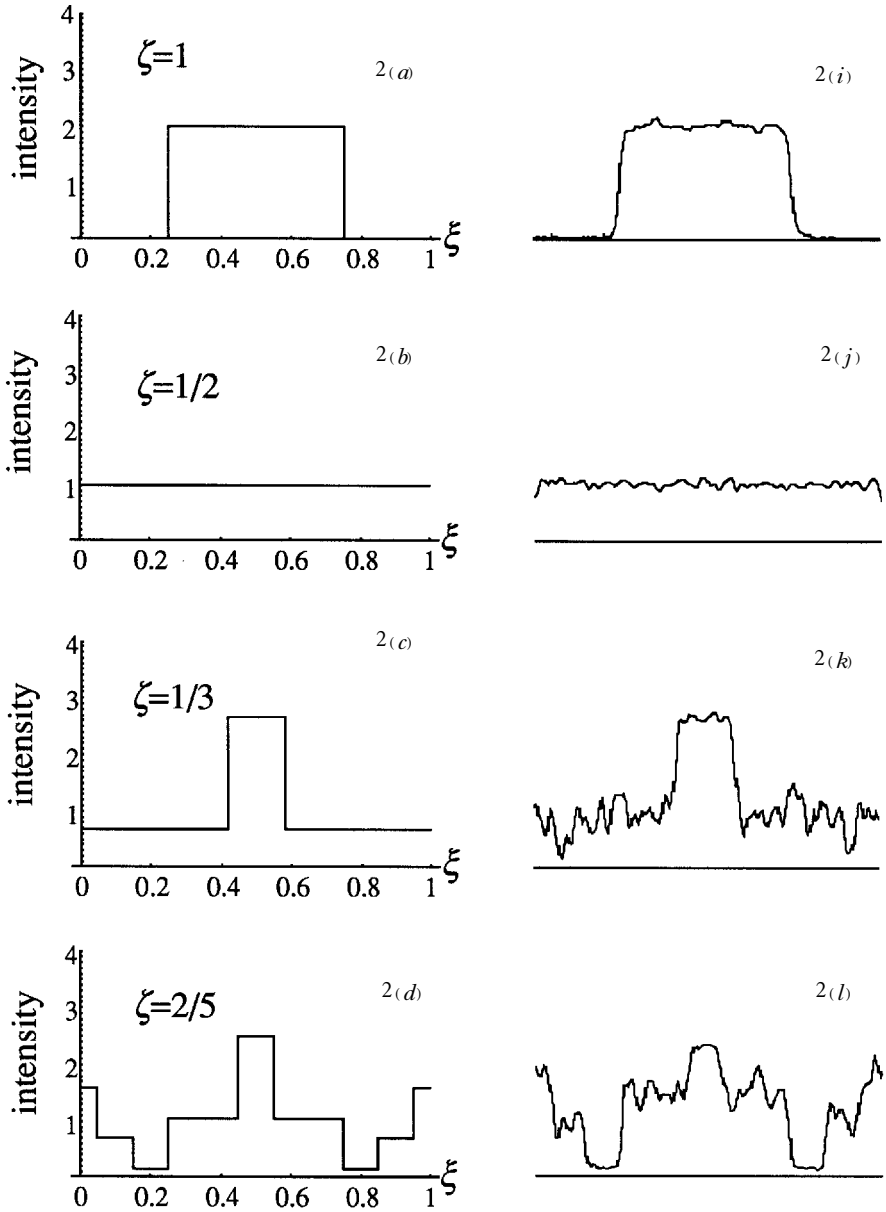
3. Fractal Talbot images

Any irrational ζ can be regarded as the limit $m \to \infty$ of a sequence of rationals p_m/q_m. According to equation (19), the corresponding Talbot images $\Psi_p(\xi, p_m/q_m)$, regarded as functions of ξ, are superpositions of q_m slit images, with phases (20). As m increases, more of these images overlap. The nature of the superposition depends on the continuity class of the grating transmission function $g(\xi)$, since this determines the decay of the Fourier coefficients g_n for large n, and hence the convergence of equation (9). Here we consider only the Ronchi grating (2), but our conclusions also hold for any amplitude and/or phase grating for which g has discontinuities, for example any grating with sharp-edged slits. Then the Talbot intensity $|\Psi_p|^2$ for irrational ζ has infinitely many infinitely small discontinuities. It is known that the result of such superpositions can be fractal [6], and we now argue that this holds in the present case.

For the Ronchi grating (2) and (7), the paraxial wave (9) is

$$\Psi_p(\xi, \zeta) = \frac{1}{2} + \frac{2}{\pi}\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \cos\left[2\pi\xi(2k+1)\right] \exp\left[-i\pi\zeta(2k+1)^2\right]. \qquad (24)$$

Considered as a periodic function of ξ with Fourier components $n = 2k + 1$, this



wave and also the intensity $|\Psi_p|^2$ have a power spectrum proportional to n^{-2}, and, for irrational ζ, pseudo-random phases $\pi\zeta n^2$. Now, it is a general result [6, 29] that the graph of a function with power spectrum $|g_n|^2$ proportional to $n^{-\beta}$ is a fractal curve with fractal dimension $D = (5 - \beta)/2$. A smooth curve has $D = 1$, and a curve with $D = 2$ is so jagged that it is almost area filling; curves with $1 \leqslant D < 2$ are continuous but non-differentiable. When applied to equation (24) the formula for D gives, for Talbot images in transverse planes specified by irrational ζ, the dimension $D_\xi = \frac{3}{2}$.

To illustrate the emergence of the transverse Talbot fractals as m increases, we

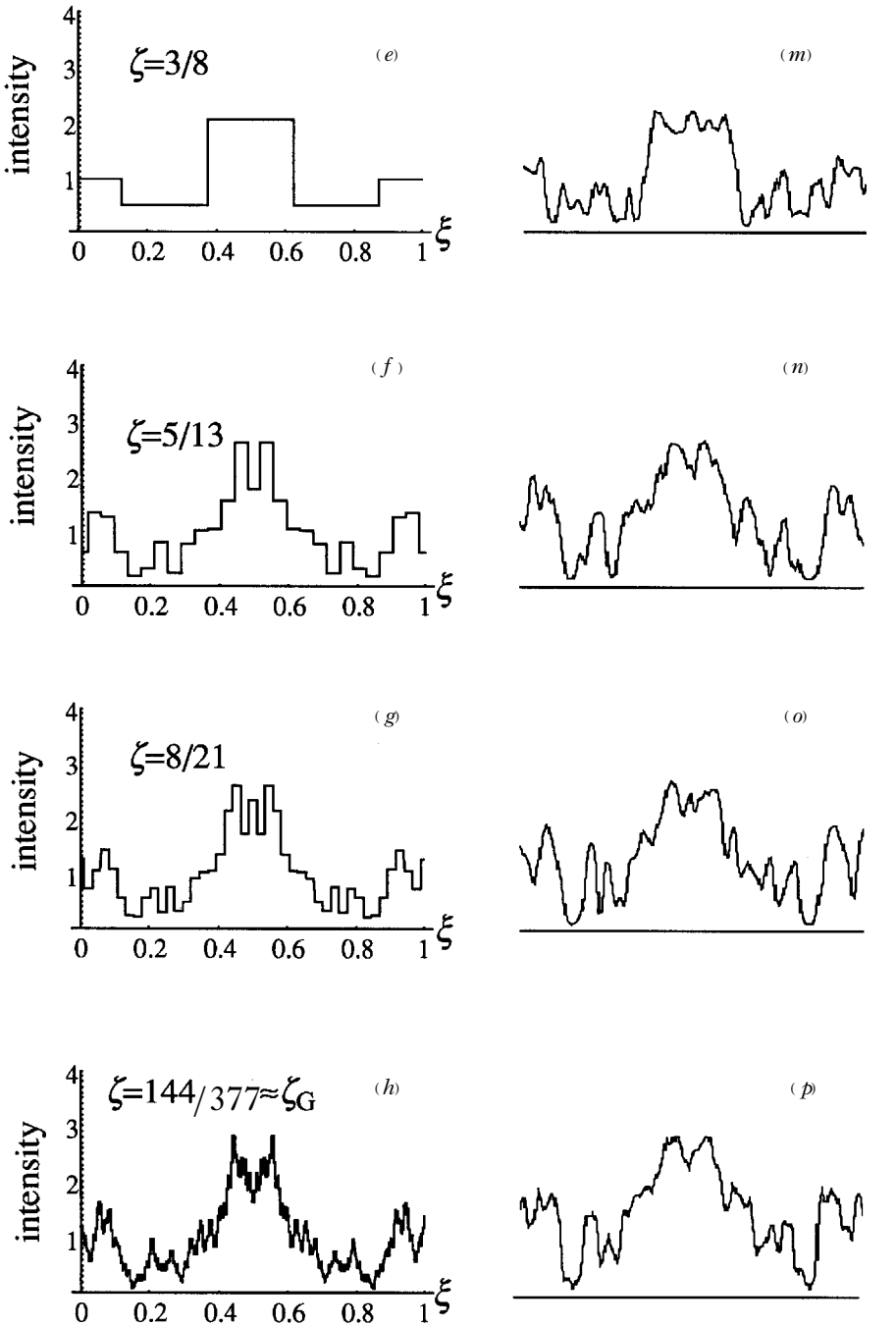


Figure 2. (*a*)–(*h*) Calculated intensities in a unit cell, for Talbot planes with $\zeta = 1$ and ζ given by the indicated rational approximations to the golden distance $\zeta_G = (3 - 5^{1/2})/2$, for the Ronchi grating (2). The transverse Talbot fractal, with dimension $D_\xi = \frac{3}{2}$, appears in (*h*). (*i*)–(*p*) Observed intensities in Talbot planes corresponding to (*a*)–(*h*), with a Ronchi grating with $a = 0{\cdot}508$ mm (50 lines in^{-1}), of which 100 lines were illuminated by a laser with $\lambda = 632{\cdot}8$ nm.

choose as the sequences p_m/q_m the successive truncations of the continued fraction [24] for ζ, namely

$$\zeta = a_0 + \frac{1}{a_1 + \left[1/(a_2 + \cdots)\right]}, \tag{25}$$

where the a_m are positive integers. This is a natural choice because these sequences give best approximations, in the sense that p_m/q_m is closer to ζ than any other fraction with denominator $q \leqslant q_m$. From this point of view the simplest irrational number is the golden mean $(5^{1/2} - 1)/2$, for which all a_m are unity, and the p_m and q_m are Fibonacci numbers. From equation (11), the Talbot images are the same for $\zeta = \zeta_G$ = 1-golden mean, that is

$$\zeta_G = \frac{3 - 5^{1/2}}{2} = 0{\cdot}381\,966\cdots = \lim\left\{0, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{2}{5}, \tfrac{3}{8}, \tfrac{5}{13}, \tfrac{8}{21}, \ldots\right\}, \tag{26}$$

and for later convenience we perform the calculations with this ζ.

Figures 2(*a*)–(*h*) show part of the sequence of Talbot images approximating that for ζ_G. The self-similarity of the fractal limit is evident in figure 2(*h*) and its magnifications given later in figures 4(*a*) and 4(*b*). Visual comparison of these figures with published curves with different fractal dimensions (for example [23]) supports the Talbot value $D = \frac{3}{2}$.

Figure 3 shows the images in two other irrational Talbot planes: $1/\pi$, which is a transcendental number, and $2^{-1/3}$, which is an algebraic number whose continued fraction is not periodic (unlike that for ζ_G and other quadratic irrationals).

Consider now the `longitudinal' Talbot intensity as a function of ζ, with transverse position ξ held constant. The wave (24) is periodic in ζ, with a Fourier series containing longitudinal frequencies restricted to the values $n = k^2$. For such a lacunary series, the power spectrum is $|g_k|^2\,\mathrm{d}k/\mathrm{d}n$, which is proportional to $n^{-3/2}$. In the argument following equation (24) we now have $\beta = \frac{3}{2}$, giving the unexpected fractal dimension $D_\zeta = (5 - \beta)/2 = \frac{7}{4}$. This is not restricted to rational ξ, and indeed we think the Talbot intensity is fractal in all longitudinal planes. The only exception occurs for the special planes $\xi = \pm\frac{1}{4}$ corresponding to the edges of the slits; in this case it follows from equation (24) that the intensity is independent of ζ.

Figure 4(*c*) shows the longitudinal Talbot fractal in the plane $\xi = 0$ corresponding to the illuminated centre of the slit; figure 4(*d*) is a magnification of the pattern, showing its self-similarity. We have calculated the patterns for several other values of ξ, and they look similar to that for $\xi = 0$.

Both sorts of fractals, and more, are contained in figure 5, which shows the `mountains of Talbot', that is the intensity $|\Psi_p|^2$ as a function of ξ and ζ, plotted as an obliquely illuminated landscape. This is a surface whose fractal dimension (obtained in [6] by adding one to the dimension of a typical section) is $\frac{11}{4}$. As well as the evident fractality almost everywhere, the original grating profile (2) is visible at $\zeta = 0$, as is the line $\xi = \frac{1}{4}$ corresponding to the special plane already mentioned.

Also visible are diagonal structures that we did not anticipate. To understand these, consider the wave (24) in the diagonal planes defined by

$$\zeta = \pm\xi + b, \tag{27}$$

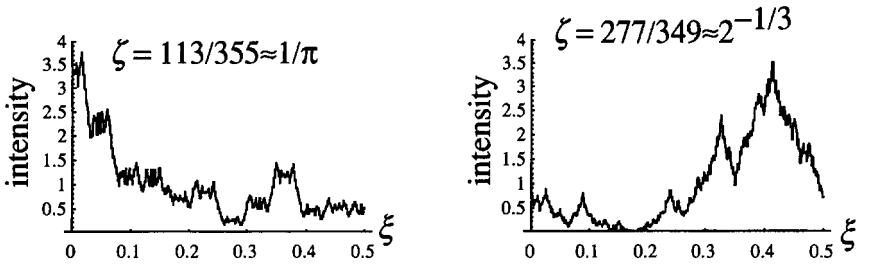


Figure 3. Transverse Talbot intensities, calculated in half-unit cells of planes approximating the distances $\zeta = 1/\pi$ and $2^{-1/3}$, for the Ronchi grating (2). Note that $1/\pi - 113/355 = 2{\cdot}7 \times 10^{-8}$, $2^{-1/3} - 277/349 = 4{\cdot}3 \times 10^{-6}$.

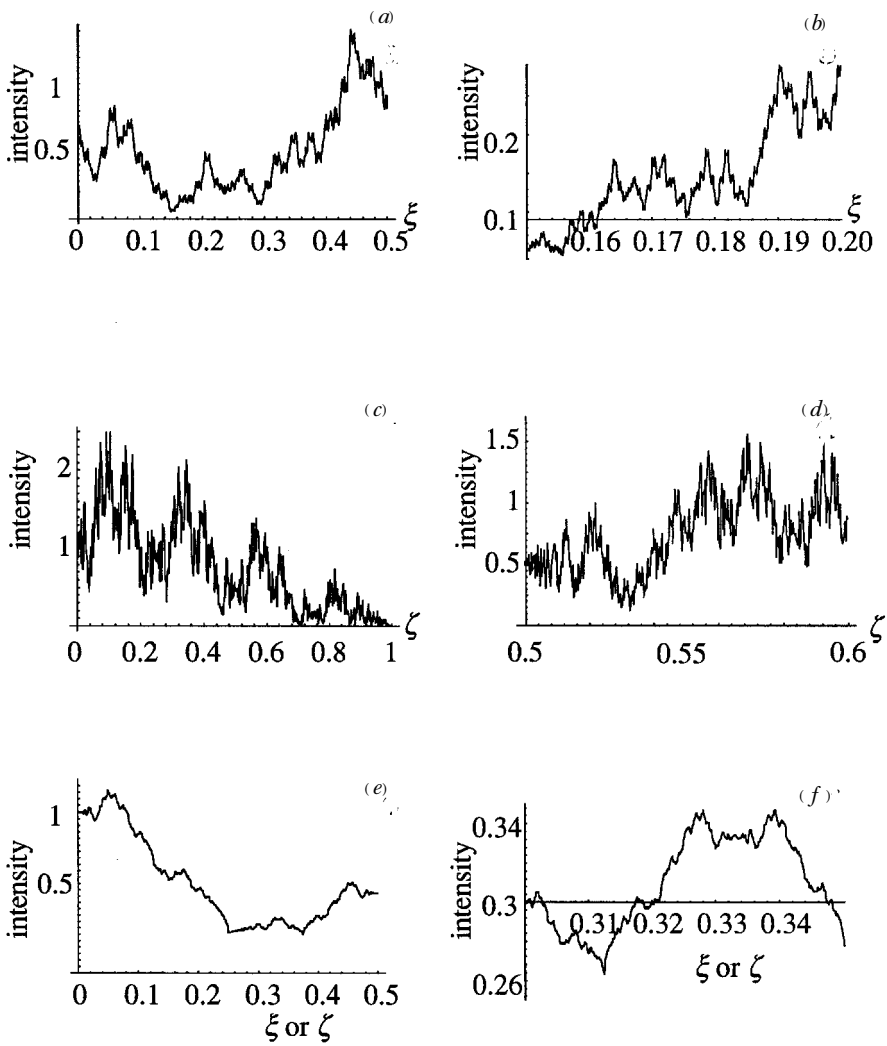


Figure 4. Talbot fractals and their magnifications: (*a*) transverse fractal, with dimension $D_\xi = \frac{3}{2}$, in the plane $\zeta = \zeta_G$; (*b*) magnification of (*a*); (*c*) longitudinal Talbot fractal, with dimension $D_\zeta = \frac{7}{4}$, in the plane $\xi = 0$; (*d*) magnification of (*c*); (*e*) diagonal Talbot fractal, with dimension $D_d = \frac{5}{4}$, in the plane $\xi = \zeta$; (*f*) magnification of (*e*).

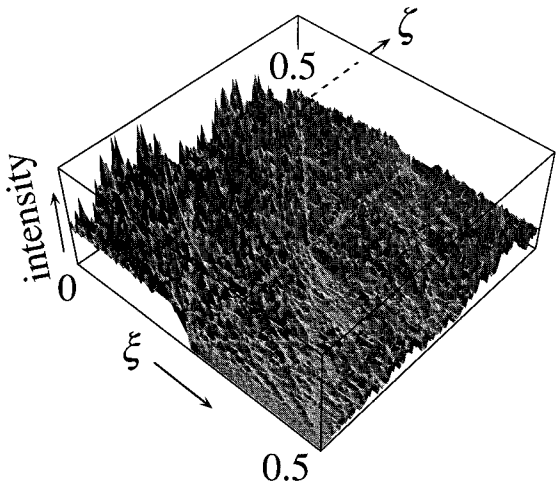


Figure 5. The `mountains of Talbot', displaying the paraxial light intensity behind a Ronchi grating as a surface with fractal dimension $\frac{11}{4}$.

where $b = 0, \pm\frac{1}{2}, \pm\frac{1}{4}$. A short analysis shows that in these planes the two exponentials in the cosines in equation (24) partially cancel for successive k, giving

$$\Psi_p(\xi, \pm\xi + b) = \frac{1}{2} - \frac{\exp\left[i\pi(-b \pm \xi)\right]}{2\pi}\left(2 + \sum_{k=0}^{\infty} \frac{(-1)^k}{(4k^2 - 1)} \exp\left(\mp i\pi\xi k^2\right)\right). \qquad (28)$$

As for longitudinal planes, this is a lacunary series with Fourier coefficients $n = k^2$, but now the intensities $|g_k|^2$ decay as k^{-4}; so the power spectrum $|g_k|^2\, dk/dn$ is proportional to $n^{5/2}$, giving $\beta = \frac{5}{2}$. It follows that in these special diagonal sections the intensity is again fractal, but with the smaller dimention $D_d = (5 - \beta)/2 = \frac{5}{4}$. One of these diagonal fractals is shown in figure 4(*e*), and again, magnified, in figure 4(*f*).

The result of our analysis of fractal aspects of the paraxial Talbot intensity, regarded as a function of ξ and ζ, can be usefully considered geometrically, as follows. Consider the space of all possible (flat) sections of the irreducible region $0 \leqslant \xi \leqslant \frac{1}{2}$, $0 \leqslant \zeta \leqslant \frac{1}{2}$ behind the grating. This space is a plane (the `plane of sections' (PS)) whose coordinates we choose to be the distance ζ_0 at which the section cuts the line $\xi = 0$, and the angle θ that it makes with the ζ direction. Each point in PS represents a section in which there is an intensity distribution whose graph is a curve G. At almost all points in PS, G has dimension $D_\zeta = \frac{7}{4}$. Along the line in PS corresponding to transverse sections, that is $\theta = \pi/2$, the G for almost all points have dimension $D_\xi = \frac{3}{2}$, exceptions being the (dense but of measure zero) set of fractional Talbot planes with rational ζ_0, whose images (section 2) have $D = 1$. Finally, at the six points in PS corresponding to diagonal sections (namely $\theta = \pi/4$ and $\zeta_0 = 0, \pm\frac{1}{4}$; $\theta = -\pi/4$ and $\zeta_0 = \frac{1}{4}, \frac{1}{2}$ and $\frac{3}{4}$), G has dimension $D_d = \frac{5}{4}$. We think the Talbot wave (24) could contain more fractals and is worth studying further. (Note added in proof: there are infinitely many diagonal sections with $D_d = 5/4$ [30].)

For almost all sections (which need not be straight) the graph of the intensity

will be the largest of the above values, namely $\frac{7}{4}$. It might appear from this that it is impossible to see the other fractals, because this would require perfect orientation of the observation plane, but in section 6 we shall find that this is not the case, and at least the transverse fractal can be detected.

As with all fractals, the finest scales of these longitudinal, transverse and diagonal Talbot fractals will be obscured by smoothing arising from physical effects neglected in the derivation given here. The two principal causes of smoothing are non-paraxiality (section 4) and the finite extent of the grating (section 5).

A surprising application of the preceding analysis, for the special case of the Ronchi grating, was pointed out to us by Dr J. H. Hannay. The solution of the paraxial wave equation in a strip waveguide, with Dirichlet boundary conditions on the sides and a constant wave amplitude across one end, is precisely the anisotropic fractal that we have found in the Talbot effect. In ref. [30] this idea is developed further. This contrasts with the solution of the corresponding problem for the heat equation, where the theta functions converge and the solution is smooth.

The preceding analysis applies whenever g has discontinuities. Gratings whose slits have edges that are not sharp would give Talbot intensities that are smooth, rather than fractal, functions of ζ and ξ.

4. Post-paraxial theory of Talbot images

To study Talbot images beyond the paraxial approximation, we first expand the square root in equation (4) to order n^4 rather than n^2. This gives the comb wave in terms of the *post-paraxial propagator* $\psi_{\mathrm{pp}}(\xi, \zeta)$, defined (cf. equation (8)) by

$$\psi_{\mathrm{comb}}(\xi, \zeta) \approx \exp(\mathrm{i}kz)\, \psi_{\mathrm{pp}}(\xi, \zeta), \tag{29}$$

where

$$\psi_{\mathrm{pp}}(\xi, \zeta) = \sum_{n=-\infty}^{\infty} \exp\left\{\mathrm{i}\pi\left[2\xi n - \zeta n^2 - \left(\frac{\lambda}{2a}\right)^2 \zeta n^4\right]\right\}.$$

It will be sufficient to specialize to the rational planes $\zeta = p/q$.

We start by following the argument of section 2 and make the substitution (12); after an application of the Poisson summation formula, this leads to the analogue of equation (15), namely

$$\psi_{\mathrm{pp}}\left(\xi\ \frac{p}{q}\right) = \frac{1}{q^{1/2}} \sum_{n=-\infty}^{\infty} A(n; q, p)\, \delta_{\mathrm{pp}}\left(\xi' - \frac{n}{q}\right), \tag{30}$$

where the post-paraxial analogue of the δ function is defined as

$$\delta_{\mathrm{pp}}(y) \equiv \int_{-\infty}^{\infty} \mathrm{d}u \exp\left\{\mathrm{i}\pi\left[2uy - \left(\frac{\lambda}{2a}\right)^2 \frac{p}{q} u^4\right]\right\}. \tag{31}$$

It is clear that the sharp Talbot images emerge in the `geometrical' limit $\lambda/a \to 0$, in which $\delta_{\mathrm{pp}}(y) \to \delta(y)$. To study the details of this emergence, it will be useful to define the *non-paraxiality parameter*

$$\epsilon \equiv \frac{\lambda}{a}\zeta^{1/2}. \tag{32}$$

We note that equation (31) is the same function as that describing diffraction through the cusp of a caustic and perpendicular to its symmetry axis [25]; the Airy function that typically arises when Gaussian integrals are generalized in analogous physical situations [11] does not occur here, because the square root in equation (4) is a function of n^2 and so its expansion does not contain n^3.

After a contour rotation, δ_{pp} can be expressed in terms of a convergent integral, as

$$\delta_{\mathrm{pp}}(y) = \frac{(4\pi)^{3/4}}{\epsilon^{1/2}} D_{\mathrm{pp}}\left(\frac{(4\pi)^{3/4}}{\epsilon^{1/2}} y\right), \tag{33}$$

where

$$D_{\mathrm{pp}}(t) \equiv \frac{\exp(-\frac{1}{8}\mathrm{i}\pi)}{\pi}\int_0^\infty \mathrm{d}x \exp(-x^4)\cos\left[xt\exp(-\tfrac{1}{8}\mathrm{i}\pi)\right]. \tag{34}$$

This is the diffraction function describing the post-paraxial blurring of the paraxially sharp Talbot images. Its principal properties are

$$\begin{aligned} &D_{\mathrm{pp}}(t) = D_{\mathrm{pp}}(-t), \qquad \int_{-\infty}^{\infty} \mathrm{d}t\, D_{\mathrm{pp}}(t) = 1, \\ &D_{\mathrm{pp}}(0) = \frac{\exp(-\frac{1}{8}\mathrm{i}\pi)}{4\pi}\Gamma(\tfrac{1}{4}), \\ &D_{\mathrm{pp}}(t) \approx \frac{\exp(-\frac{1}{4}\mathrm{i}\pi)}{2^{5/6}(3\pi)^{1/2}|t|^{1/3}}\exp\left[\frac{3}{4}\mathrm{i}\left(\frac{t^4}{4}\right)^{1/3}\right] \quad (|t| \gg 1). \end{aligned} \tag{35}$$

Figure 6 shows graphs of the real and imaginary parts of $D_{\mathrm{pp}}(t)$.

For the Ronchi grating (2), the post-paraxial replacement $g_{\mathrm{pp}}(\xi\,\epsilon)$ of the paraxially perfect reproductions $g(\xi)$ of the individual slits is obtained from equations (3), (29), (30), (33) and (34) as

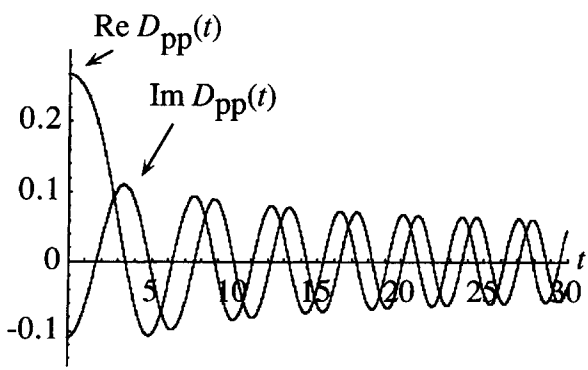


Figure 6. Post-paraxial generalization (34) of the δ function, describing the blurring of the Talbot images.

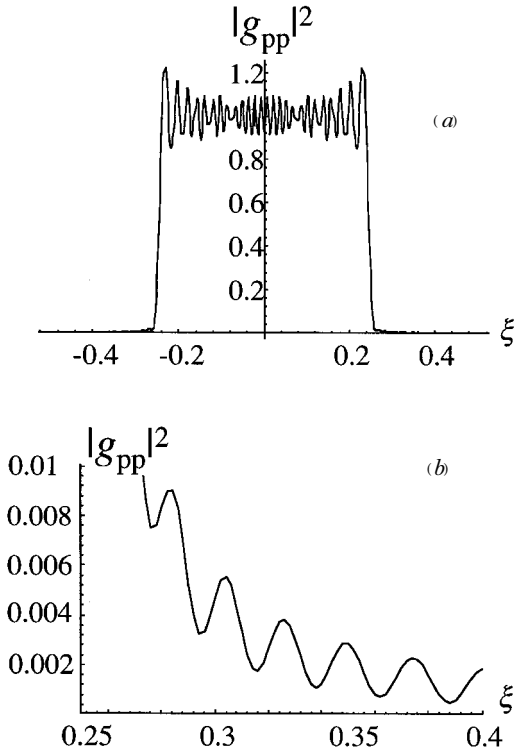


Figure 7. (*a*) Post-paraxial intensity (36) of the Talbot image of a slit of a Ronchi grating with $\epsilon = 0{\cdot}001\,246$; this corresponds, for example, to the Talbot plane $\zeta = 1$ of a grating with $a = 0{\cdot}508$ mm (50 lines in^{-1}), illuminated by light with $\lambda = 632{\cdot}8$ nm. (*b*) Magnification of (*a*) showing weak fringes on the dark side of the slit.

$$g_{pp}(\xi, \epsilon) = \frac{1}{\pi}\int_0^\infty dx \exp(-x^4)\frac{\sin\left[c(\epsilon)x(\xi+\tfrac{1}{4})\right] - \sin\left[c(\epsilon)x(\xi-\tfrac{1}{4})\right]}{x}, \qquad (36)$$

where

$$c(\epsilon) \equiv \frac{(4\pi)^{3/4}}{\epsilon^{1/2}} \exp(-\tfrac{1}{8}i\pi).$$

This is the difference between two integrals of the functions D_{pp}, corresponding to post-paraxial diffraction from the two edges of the slits. An example of the blurring introduced by this violation of paraxiality is shown in figure 7.

A conclusion from this section is that the sharp edges of the Talbot images are replaced post-paraxially by fringes whose size, from the scalings (32) and (33) and the observation (from figure 6) that the first zero of $\mathrm{Re}\left[D_{pp}(t)\right]$ is near $t = \pi$ is

$$\Delta\xi_{pp,edge} = \frac{\pi^{1/4}}{2^{3/2}}\epsilon^{1/2} = 0{\cdot}47\zeta^{1/4}\left(\frac{\lambda}{a}\right)^{1/2}. \qquad (37)$$

Thus the degradation of the Talbot images gets slowly worse with increasing distance from the grating. It is interesting to compare $\Delta\xi_{pp,edge}$ with the size $\Delta\xi_{edge}$

of the fringes associated with ordinary Fresnel (i.e. paraxial) diffraction from the edge of a single slit. Standard theory [26] gives

$$\Delta\xi_{\text{edge}} = \zeta^{1/2}. \tag{38}$$

In the Talbot effect, coherent diffraction from the infinite array of slits reduces the scale of this blurring to zero paraxially, and to $\Delta\xi_{\text{pp,edge}}$ post-paraxially. The reduction factor is

$$\frac{\Delta\xi_{\text{pp,edge}}}{\Delta\xi_{\text{edge}}} = \frac{\pi^{1/4}}{2^{3/2}}\left(\frac{\lambda}{z}\right)^{1/4}, \tag{39}$$

which is very small in all contexts relevant to the Talbot effect.

The analytical theory presented here, in which the square root in equation (4) is expanded to order n^4, gives the lowest-order post-paraxial correction to the Fresnel approximation. It is of course possible to make exact computations of the blurring of the edges of the slits in the Talbot images. However, our numerical explorations, and comparison with some published computations [7], show that the theory based on the function (34) gives results indistinguishable from exact computations for all gratings with $\lambda/a \ll 1$ that are interesting for the Talbot effect.

Nevertheless we can use the full nonparaxial theory, in the form (5), to give two other measures of the sensitivity of the Talbot images (in ζ as well as ξ). These are based on the observation that evanescent waves $|n| > a/\lambda$ can be neglected in the sum (5); even the first such wave is of order $\exp\left[-\zeta(a/\lambda)^{3/2}\right]$.

With the sum cut off at $|n| = n_{\max} = \text{int}\,(a/\lambda)$, the fastest possible ξ and ζ variations come from $n = \pm n_{\max}$ and $n = 0$ respectively, giving

$$\Delta\xi_{\text{pp,min}} = \frac{\lambda}{a}, \qquad \Delta\zeta_{\text{pp,min}} = \left(\frac{\lambda}{a}\right)^2. \tag{40}$$

When converted to variations in x and z, these blurring distances simply reflect the fact that no non-evanescent monochromatic wave field can vary on scales smaller than λ. However, the decay of $|g_n|$ means that these rapid variations are extremely weak, and equation (40) gives unrealistic measures of the non-paraxial blurring.

More realistic are the measures of ξ and ζ sensitivity calculated by spatial averaging of the intensity and its derivatives. We shall present general formulae and evaluate them explicitly for the Ronchi grating (7). Defining

$$d_n = \left(\frac{a}{\lambda}\right)^2\left[1 - \left(\frac{n\lambda}{a}\right)^2\right]^{1/2} \tag{41}$$

and, neglecting evanescent waves, we obtain, from equation (5), the intensity as

$$I(\xi,\zeta) \equiv \left|\Psi(\xi,\zeta)\right|^2 = \sum_{|m,n|<a/\lambda}\sum g_m^* g_n \exp\left\{2i\pi\left[\xi(n-m) + \zeta(d_n - d_m)\right]\right\}. \tag{42}$$

To obtain the mean intensity, we average over a period of the Talbot plane, that is over $0 \leqslant \xi \leqslant 1$, thereby eliminating the terms with $m \neq n$:

$$\langle I \rangle = \sum_{|n|<a/\lambda} |g_n|^2. \tag{43}$$

Evaluating the sum for the Ronchi grating (appendix) gives, to leading order in λ/a,

$$\langle I \rangle \approx \frac{1}{2} - \frac{\lambda}{\pi^2 a}. \tag{44}$$

To obtain an estimate of the fluctuations, we study the average

$$\langle I^2 \rangle = \left\langle \sum\sum\sum\sum_{|m,n,r,s|<a/\lambda} g_m^* g_n g_s g_r^* \exp\left\{2i\pi\left[\xi(n - m - r + s) + \zeta(d_n - d_m - d_r + d_s)\right]\right\} \right\rangle. \tag{45}$$

Now the average is over a region of space, that is over both a grating period $0 \leqslant \xi \leqslant 1$ in each Talbot plane and a sufficiently large interval in ζ to eliminate all cross-phases d_n; the surviving terms are those where $n = m$ and $r = s$, and where $n = r$ and $m = s$. Thus

$$\langle I^2 \rangle = 2\sum\sum_{|m,n|<a/\lambda} |g_n|^2 |g_m|^2 = 2\langle I \rangle^2. \tag{46}$$

This shows that on the average the fluctuations of the Talbot wave mimic those of a function with random phases and power spectrum $|g_n|^2$.

The simplest average describing intensity variations within a Talbot plane is

$$I_\xi \equiv \left\langle \left(\frac{\partial I}{\partial \xi} \right)^2 \right\rangle^{1/2}. \tag{47}$$

In terms of this quantity, we can define an average blurring $\Delta\xi_{\mathrm{pp,av}}$ of a Talbot image as the distance over which the fractional change of intensity is unity. Thus

$$\Delta\xi_{\mathrm{pp,av}} \equiv \frac{\langle I \rangle}{I_\xi}. \tag{48}$$

To get an expression for I_ξ, we differentiate equation (42) and use the same argument that led from equation (45) to equation (46):

$$I_\xi^2 = 4\pi^2 \sum\sum_{|r,s|<a/\lambda} (r - s)^2 |g_r|^2 |g_s|^2. \tag{49}$$

For the Ronchi grating, evaluation of the sums (appendix) leads to

$$\Delta\xi_{\mathrm{pp,av}} = \frac{1}{4}\left(\frac{\lambda}{a}\right)^{1/2}. \tag{50}$$

Comparing this result with $\Delta\xi_{\mathrm{pp,edge}}$ (equation (37)), we note the absence of the

factor $\zeta^{1/4}$, indicating that at many Talbot distances the post-paraxial blurring of the edges of the slit images exceeds the blurring elsewhere in the images.

In an analogous manner, the intensity variations between Talbot planes can be described by

$$I_\zeta \equiv \left\langle \left(\frac{\partial I}{\partial \zeta} \right)^2 \right\rangle^{1/2}. \tag{51}$$

Then, an average `depth of focus' $\Delta\zeta_{\mathrm{p,av}}$ of the Talbot images can be defined by (cf. equation (48))

$$\Delta\zeta_{\mathrm{pp,av}} \equiv \frac{\langle I \rangle}{I_\zeta}. \tag{52}$$

The analogue of equation (49) is

$$I_\zeta^2 = \pi^2 \sum_{|r,s|<a/\lambda} \sum (d_r - d_s)^2 |g_r|^2 |g_s|^2 \tag{53}$$

and leads, for the Ronchi grating (appendix), to

$$\Delta\zeta_{\mathrm{pp,av}} = \frac{1}{4(\pi - 3)^{1/2}} \left(\frac{\lambda}{a} \right)^{3/2} = 0{\cdot}66 \left(\frac{\lambda}{a} \right)^{3/2}. \tag{54}$$

5. Finitely many slits

To sufficient approximation, the paraxial wave $\Psi_{\mathrm{p},N}(\xi, \zeta)$ for a finite grating with N slits is obtained from the representation (9), for the infinite grating, by truncating the sum over diffracted beams, each regarded as having width Na, to include those that overlap at the point (x, z). An elementary geometrical argument [8] shows that near the middle of the Talbot images, that is for $|\xi| \ll N/2$, the contributing diffracted beams are those with $|n| < N/2\zeta$ (the limit of `walk-off' of the diffracted beams). Extensive numerical and analytical examination of this approximation [8] shows that it is very accurate when $N \gg 1$. Thus

$$\Psi_{\mathrm{p},N}(\xi, \zeta) = \sum_{n=-n_N}^{n_N} g_n \exp(2\pi \mathrm{i} \xi n - \mathrm{i}\pi\zeta n^2) \quad (\xi \ll \tfrac{1}{2}N), \tag{55}$$

where $n_N \equiv \mathrm{int}\,(N/2\zeta)$.

To cast this in a form that exhibits the Talbot effect, we consider the wave from a finite comb of δ slits, by taking $g_n = 1$ in equation (55), restrict ourselves to the rational planes $\zeta = p/q$ and split the sum as in equation (12), where now l runs between $\pm n_N/q$. Thus the *N-slit paraxial propagator* becomes

$$\psi_{\mathrm{p},N}\left(\xi, \frac{p}{q}\right) = \sum_{s=1}^{q} \exp\left[\mathrm{i}\pi \left(2s\xi - \frac{p}{q} s^2 \right) \right] \frac{\sin\left[\pi\xi'(2n_N + q)\right]}{\sin(\pi q \xi')}, \tag{56}$$

where ξ' is defined by equation (16). For finite N, the quantity $\sin\left[\pi\xi'(2n_N + q)\right]/\sin(\pi q\xi')$ describes the blurring of the Talbot images (within

the paraxial approximation) and so is the (unsurprising) finite-slit analogue of the post-paraxial generalization $D_{\mathrm{pp}}(t)$ (equation (34)) of the δ function. In the limit $N \to \infty$ of an infinite grating, the quantity $\sin\left[\pi\xi'(2n_N + q)\right]/\sin(\pi q\xi')$ tends to the Talbot δ comb, that is

$$\frac{\sin\left[\pi\xi'(2n_N + q)\right]}{\sin(\pi q\xi')} \to \frac{1}{q}\sum_{n=-\infty}^{\infty}\delta\left(\xi' - \frac{n}{q}\right) \quad \text{as } N \to \infty \tag{57}$$

and the sum over s, when evaluated at the positions ξ of the Talbot images, gives the phase factors $A(n; q, p)$, so that the paraxial propagator (15) for the infinite grating is regained.

For large N, the quantity $\sin\left[\pi\xi'(2n_N + q)\right]/\sin(\pi q\xi')$ in equation (56) is dominated by peaks at the Talbot positions $\xi' = n/q$, with subsidiary narrow fringes whose widths give a measure of the blurring of the Talbot images. The widths are

$$\Delta\xi_{N,\min} = \frac{\zeta}{N}. \tag{58}$$

We use the subscript min because this blurring is precisely that corresponding to the highest Fourier component in equation (55). A more thorough exploitation of this idea [20] includes the degradation of Talbot images outside the central region $\xi \approx 0$. The same Fourier component gives the fastest ζ variation, and hence the `depth of focus'

$$\Delta\zeta_{N,\min} = \frac{4\zeta^2}{N^2} \tag{59}$$

(see also [27]).

Now we study the spatially averaged ξ and ζ sensitivity of the finite-N paraxial intensity, following the analogous treatment, in the previous section, of the non-paraxial fluctuations. The intensity is

$$\begin{aligned} I_N(\xi, \zeta) &\equiv \left|\Psi_{\mathrm{p},N}(\xi, \zeta)\right|^2 \\ &= \sum_{|m,n| \leqslant n_N}\sum g_m^* g_n \exp\left\{\mathrm{i}\pi\left[2\xi(n - m) - \zeta(n^2 - m^2)\right]\right\}. \end{aligned} \tag{60}$$

By analogy with equations (47), (48), (51) and (52) we define the mean square intensity derivatives, and ξ and ζ blurrings, by

$$\begin{aligned} I_{N,\xi} &\equiv \left\langle\left(\frac{\partial I_N}{\partial\xi}\right)^2\right\rangle^{1/2}, & I_{N,\zeta} &\equiv \left\langle\left(\frac{\partial I_N}{\partial\zeta}\right)^2\right\rangle^{1/2}, \\ \Delta\xi_{N,\mathrm{av}} &\equiv \frac{\langle I_N\rangle}{I_{N,\xi}}, & \Delta\zeta_{N,\mathrm{av}} &\equiv \frac{\langle I_N\rangle}{I_{N,\zeta}}. \end{aligned} \tag{61}$$

The derivatives are given by equations (49) and (53) with the limits in the sum replaced by $\pm n_N$ and d_n replaced by $-n^2/2$. For the Ronchi grating (7), the average image blurring and depth of focus are

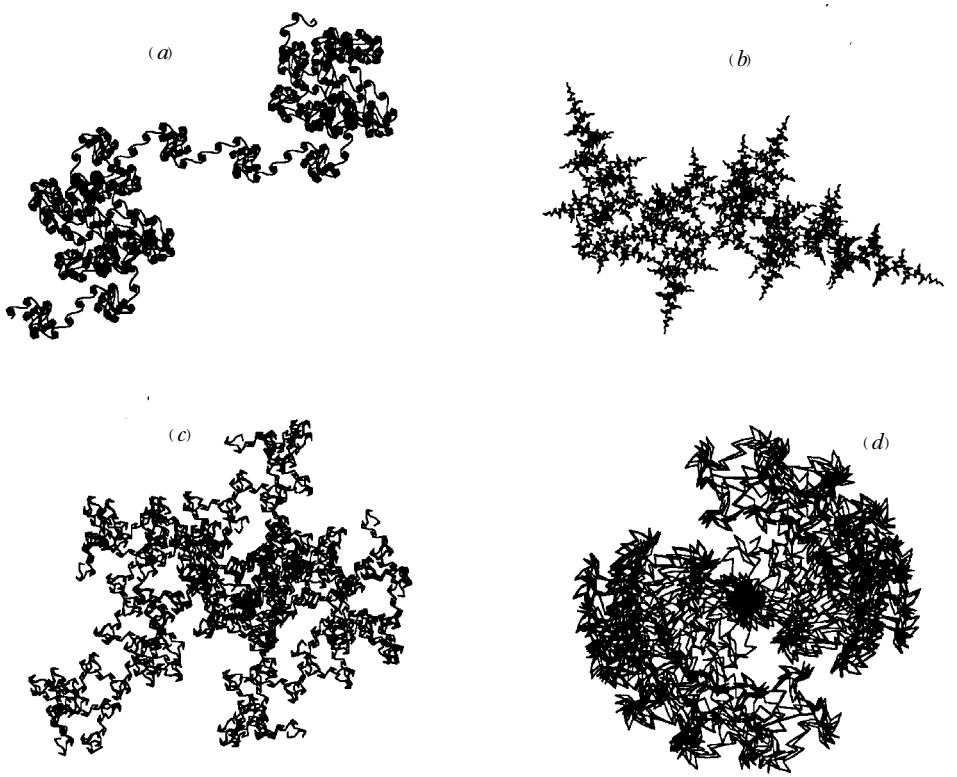


Figure 8. `Curlicues' traced in the complex plane with axes Re S, Im S by the Talbot sums (63) for $1 \leqslant L \leqslant 5000$, for (*a*) $\zeta = 50^{1/2} - 7 \approx 0{\cdot}071\,07$, (*b*) $\zeta = (5^{1/2} - 1/2 \approx 0{\cdot}618\,03$, (*c*) $\zeta = 2^{-1/3} \approx 0{\cdot}793\,70$ and (*d*) $\zeta = 1/\pi$.

$$\Delta\xi_{N,\mathrm{av}} = \left(\frac{\zeta}{8N}\right)^{1/2}, \qquad \Delta\zeta_{N,\mathrm{av}} \equiv 6^{1/2}\left(\frac{\zeta}{N}\right)^{3/2} \tag{62}$$

(the sum involved in $\Delta\zeta_{N,\mathrm{av}}$ is evaluated in the appendix).

With finite N, there is a transition from the Fresnel to the Fraunhofer regime. We do not pursue this, because it has been studied in detail [28] elsewhere.

It is very interesting to study how the Talbot image wave at a given field point (ξ, ζ) varies as the number N of grating slits is increased. The dependence is qualitatively the same for any ξ and for convenience we choose the centre $\xi = 0$. Then for narrow slits, the wave (55) (with $g_n = 1$) depends on the sum

$$S_L(\zeta) = \sum_{n=1}^{L} \exp(\mathrm{i}\pi\zeta n^2) \quad (L \gg 1). \tag{63}$$

This class of sums has been analysed in detail [8]. As L increases, the sum of unit vectors in the complex plane of $S_L(\zeta)$, representing the phase and intensity of the Talbot wave, traces out curves of fantastic complexity, whose structural elements, `curlicues', are discrete generalizations of the Cornu spiral. The curves have a

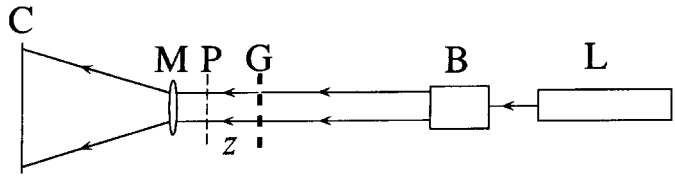


Figure 9. Experiment to observe Talbot images. Light from a 10 mW He–Ne laser L (beam diameter, 1 mm) passes through a 50× beam broadener B and strikes a moveable Ronchi grating G. The image in the desired Talbot plane P, at distance $z = \zeta z_T$ from G, is magnified by a 20× microscope objective M and focused onto the film plane of a lensless camera C.

hierarchical structure that depends sensitively on the arithmetic nature of the Talbot position ζ, as figure 8 illustrates. To understand the hierarchy, a renormalization transformation of $S_L(\zeta)$ was developed [8], in which the sum for given L and ζ is related to a magnified version of the sum with a smaller value of L and a different value ζ', related to ζ by a chaotic mapping. The same mapping, applied directly to the paraxial propagator (8) (i.e. for infinite N), will probably yield further insight into the fractal `mountains of Talbot' (figure 5), but we do not pursue this aspect here.

6. Experiments with a Ronchi grating

Using laser light with $\lambda = 632{\cdot}8$ nm (figure 9), we illuminated a Ronchi grating with 50 lines in^{-1}, that is with $a = 0{\cdot}508$ mm. Therefore the Talbot distance was $z_T = 407{\cdot}81$ mm. This means that the natural `unit cell' of the Talbot effect, namely $\Delta x = a$, $\Delta z = z_T$, was enormously elongated (by a factor $a/\lambda = 803$) relative to the dimensionless unit cell $\Delta\xi = 1$, $\Delta\zeta = 1$. The laser beam was broadened and spatially filtered into a parallel beam with a diameter of 2 in; so the number of illuminated grating periods was $N = 100$. The aim of the experiments was to record the transverse and longitudinal fractal intensity patterns described in section 2.

Consider first the transverse fractal, with dimension $D_\xi = \frac{3}{2}$. In the ideal case of perfect paraxiality and infinite N, observation of this fractal would require perfect orientation of the screen, in order to avoid seeing the typical fractal with $D_\zeta = \frac{7}{4}$. In reality, observation is possible because of the blurring and defocusing described in sections 4 and 5. Moreover, it is greatly facilitated by the elongation of the pattern in the z direction, because a theoretically allowed misorientation of the screen from perfect transversality in $\xi\,\zeta$ space is multiplied by a/λ in the physical x, z space.

We magnified the Talbot images (figure 9), and photographed them in a series of transverse planes near $\zeta = \zeta_G =$ 1-golden mean $= (3 - 5^{1/2})/2 = 0{\cdot}381\,966$. Each image was scanned into a computer, and the intensity profiles within a unit cell were determined by analysing the scans with a densitometer program. These experimental profiles are shown in figures 2(*i*)–(*p*). They should be compared with the corresponding theoretical intensity profiles in figures 2(*a*)–(*h*). Perfect agreement is not expected, because of the very blurring and defocusing that makes the observation possible. Before estimating these effects, we first note that the non-paraxiality parameter (32) was $\epsilon = 0{\cdot}000\,77$.

In our experiment, the different measures of defocusing defined by equations (40), (54), (59) and (62) were

$$\begin{aligned} \Delta\zeta_{\mathrm{pp,min}} &= 3{\cdot}15\times10^{-6}, & \Delta\zeta_{\mathrm{pp,av}} &= 2{\cdot}9\times10^{-5}, \\ \Delta\zeta_{N,\mathrm{min}} &= 5{\cdot}8\times10^{-5}, & \Delta\zeta_{N,\mathrm{av}} &= 5{\cdot}8\times10^{-4}. \end{aligned} \tag{64}$$

The largest, and therefore theoretically dominating, of these defocusings is that corresponding to the average intensity variation due to the finite number of slits; it corresponds to a defocusing distance $z_{\mathrm{T}}\Delta\zeta_{N,\mathrm{av}} = 0{\cdot}24\,\mathrm{mm}$. This maximum defocusing limits the degree of rational approximation to ζ_{G} that we can hope to achieve. Because

$$\begin{aligned} &\zeta_{\mathrm{G}} - \tfrac{1}{2} = -0{\cdot}118, && \zeta_{\mathrm{G}} - \tfrac{1}{3} = 0{\cdot}049, && \zeta_{\mathrm{G}} - \tfrac{2}{5} = -0{\cdot}018, \\ &\zeta_{\mathrm{G}} - \tfrac{3}{8} = 0{\cdot}0070, && \zeta_{\mathrm{G}} - \tfrac{5}{13} = -0{\cdot}0026, && \zeta_{\mathrm{G}} - \tfrac{8}{21} = 0{\cdot}0010, \\ &\zeta_{\mathrm{G}} - \tfrac{13}{34} = -3{\cdot}9\times10^{-4}, && \zeta_{\mathrm{G}} - \tfrac{144}{377} = 3{\cdot}15\times10^{-6}, \end{aligned} \tag{65}$$

we can hope to distinguish $\zeta = \frac{8}{21}$, but not $\zeta = \frac{13}{34}$, from ζ_{G}.

Whether in fact these approaches to the transverse fractal Talbot image can be distinguished depends on the blurring within the Talbot planes. The different measures of blurring defined by equations (37), (40), (50), (58) and (62) were, in our experiment,

$$\begin{aligned} \Delta\xi_{\mathrm{pp,edge}} &= 0{\cdot}0130, & \Delta\xi_{\mathrm{pp,min}} &= 0{\cdot}0012, & \Delta\xi_{\mathrm{pp,av}} &= 0{\cdot}0088, \\ \Delta\xi_{N,\mathrm{min}} &= 0{\cdot}0038, & \Delta\xi_{N,\mathrm{av}} &= 0{\cdot}022. \end{aligned} \tag{66}$$

The largest, and therefore theoretically dominating, of these blurrings is again that corresponding to the average intensity variation due to the finite number of slits; it corresponds to a blurring distance $a\Delta\xi_{N,\mathrm{av}} = 11\,\mu\mathrm{m}$ (in order to photograph these images, they were magnified, and this blurring distance became much greater than the grain size of the film).

These theoretical estimates show that in our experiment the resolution is limited by the number of slits rather than paraxiality. They suggest that we can hope to distinguish detail in the transverse Talbot fractal down to about 2% of the grating spacing. Inspection of the fine structure of figure 2 shows that we do in fact achieve this; to the expected resolution, the positions and, roughly, the heights, of the observed intensity minima and maxima correspond to those expected theoretically. (We did not attempt to measure the fractal dimension of the curve in figure 2(*p*) (e.g. by box counting), and so to test the theoretical prediction of $D_\xi = \frac{3}{2}$, because such measurements are notoriously unreliable for small data sets such as that in our experiment.)

Now consider the longitudinal Talbot fractal, with dimension $D_\zeta = \frac{7}{4}$. In principle it is possible to measure this by photographing the image from a tilted grating, but in practice the resolution is too low. Therefore we modified the arrangement in figure 9 by replacing the camera with a photodiode and measuring the intensity directly. Careful alignment was necessary, because the previously mentioned elongation of the pattern in the z direction greatly magnifies any misorientation (in contrast with the transverse case, where misorientation is reduced) and introduces the risk of getting lost in the `mountains of Talbot',

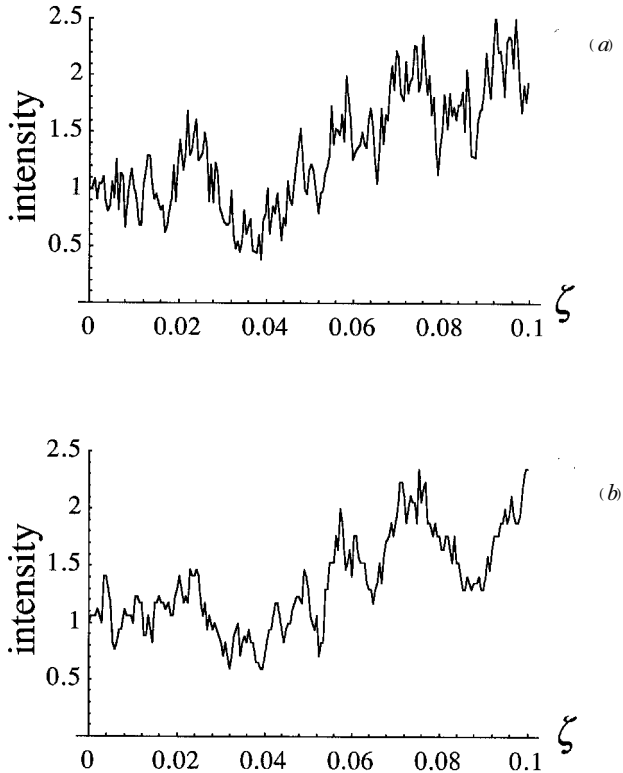


Figure 10. Longitudinal Talbot fractal, with dimension $D_\zeta = \frac{7}{4}$, in the plane $\xi = 0$, for the range $0 \leqslant \zeta \leqslant 0{\cdot}1$: (*a*) computed from equation (24) and sampled at 200 points (corresponding to intervals $\Delta\zeta = 5 \times 10^{-4}$); (*b*) measured at intervals $\Delta z = 0{\cdot}2$ mm (also corresponding to $\Delta\zeta = 5 \times 10^{-4}$).

which makes comparison with theory difficult. Therefore we exploited the symmetry of the pattern and measured the intensity for $\xi = 0$. Our measurements covered a z range of 40 mm, corresponding to a ζ range of 0·1. Over this range, it was necessary to keep the fringe pattern centred on the pinhole of the photodiode to an absolute accuracy of 5 μm, because, as explained following equation (66), the expected blurring of the fractals in x was about 10 μm. We measured the intensity at intervals in z of 0·2 mm, slightly less than the expected blurring of 0·24 mm explained after equation (64).

The result of these measurements is shown in figure 10(*b*), to be compared with the theory in figure 10(*a*) which is calculated with the same number of sample points. Although the observed blurring seems slightly greater than that estimated theoretically, the agreement is convincing and demonstrates the existence of the longitudinal fractal.

For historical interest, we repeated Lord Rayleigh's [3] experimental study of the Talbot effect in white light. We used a halogen bulb instead of his gas flame, a green interference filter ($\Delta\lambda/\lambda \approx 0{\cdot}1$) instead of his coloured glass filter

($\Delta\lambda/\lambda \approx 0{\cdot}5$) and, following his recommendation, a collimating lens. The results (not shown here) are crude in comparison with those of our laser experiments. The transverse Talbot fractals are very blurred, and Rayleigh, with his coloured glass, could have seen only the most prominent features.

Appendix. Fluctuation sums for the Ronchi grating

For the grating (7), the mean intensity (43) is

$$\begin{aligned}\langle I\rangle &= \frac{1}{4}+\frac{1}{\pi^2}\sum_{|k|<a/2\lambda}\frac{1}{(2k+1)^2}\\ &= \frac{1}{4}+\frac{2}{\pi^2}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}-\frac{2}{\pi^2}\sum_{k=a/2\lambda}^{\infty}\frac{1}{(2k+1)^2}.\end{aligned} \tag{A 1}$$

The penultimate sum is $\pi^2/8$, and the last sum can be approximated by an integral, giving equation (44).

For the ξ derivative (49), we have

$$I_\xi^2 = 8\pi^2\sum_{|r<a/\lambda|} r^2|g_r|^2\sum_{|s<a/\lambda|}|g_s|^2 = 8\sum_{|k<a/2\lambda|}1\times\frac{1}{2} = 4\frac{a}{\lambda}, \tag{A 2}$$

leading directly to equation (50).

For the ζ derivative (53), we have

$$I_\xi^2 = 2\pi^2\left[\sum_{|r|<a/\lambda} d_r^2|g_r|^2\sum_{|s|<a/\lambda}|g_s|^2-\left(\sum_{|r|<a/\lambda} d_r|g_r|^2\right)^2\right]. \tag{A 3}$$

This involves three different sums. The first is (using equation (41))

$$\begin{aligned}\sum_{|r|<a/\lambda} d_r^2|g_r|^2 &= \left(\frac{a}{\lambda}\right)^4\left[\sum_{|r|<a/\lambda}|g_r|^2-\left(\frac{\lambda}{a}\right)^2\sum_{|r|<a/\lambda} r^2|g_r|^2\right]\\ &\approx \left(\frac{a}{\lambda}\right)^4\left[\frac{1}{2}-\frac{\lambda}{\pi^2 a}-\left(\frac{\lambda}{\pi a}\right)^2\sum_{|k|<a/2\lambda}1\right]\\ &\approx \left(\frac{a}{\lambda}\right)^4\left(\frac{1}{2}-\frac{2\lambda}{\pi^2 a}\right).\end{aligned} \tag{A 4}$$

The second sum is just equation (44). The third sum is

$$\sum_{|r|<a/\lambda} d_r|g_r|^2 = \left(\frac{a}{\lambda}\right)^2 \sum_{|r|<a/\lambda} |g_r|^2 \left\{1 + \left[1 - \left(\frac{r\lambda}{a}\right)^2\right]^{1/2} - 1\right\}$$

$$\approx \left(\frac{a}{\lambda}\right)^2 \left(\frac{1}{2} - \frac{\lambda}{\pi^2 a} + \frac{2}{\pi^2}\int_0^{a/2\lambda} \mathrm{d}k \frac{\left[1 - (2k\lambda/a)\right]^{1/2} - 1}{4k^2}\right)$$

$$= \frac{1}{2}\left(\frac{a}{\lambda}\right)^2 \left(1 - \frac{\lambda}{\pi a}\right). \tag{A 5}$$

On combining these, equation (A 3) becomes

$$I_\zeta^2 \approx 4\left(\frac{a}{\lambda}\right)^3 (\pi - 3), \tag{A 6}$$

from which equation (54) follows immediately.

Finally, the second result in equation (62) depends on equation (61) and

$$I_{N,\zeta}^2 = \pi^2 \sum_{|r,s|<N/2\zeta} \sum (r^2 - s^2)^2 |g_r|^2 |g_s|^2$$

$$= 2\pi^2 \left[\sum_{|r|<N/2\zeta} r^4 |g_r|^2 \sum_{|s|<N/2\zeta} |g_s|^2 - \left(\sum_{|r|<N/2\zeta} r^2 |g_r|^2\right)^2\right]$$

$$= 2\left[\sum_{|k|<N/4\zeta} (2k+1)^2 \frac{1}{2} - \frac{1}{\pi^2}\left(\sum_{|k|<N/4\zeta} 1\right)^2\right]$$

$$\approx \frac{1}{24}\left(\frac{N}{\zeta}\right)^3. \tag{A 7}$$

Here the last equality relies on the fact that in the penultimate member the first sum dominates and can be replaced by an integral. The second result in equation (62) follows at once.

Acknowledgments

We thank Professor W. Dultz for providing experimental equipment, and Dr. J. P. Keating for a helpful suggestion.

References

[1] Talbot, H. F., 1836, *Phil. Mag.*, **9,** 401–407.
[2] Patorski, K., 1989, *Prog. Optics*, **27,** 1-108.
[3] Rayleigh, Lord, 1881, *Phil. Mag.*, **11,** 196–205.
[4] Hiedemann, E. A., and Breazale, M. A., 1959, *J. opt. Soc. Am.*, **49,** 372–375.
[5] Winthrop, J. T., and Worthington, C. R., 1965, *J. opt. Soc. Am.*, **55,** 373–381.
[6] Mandelbrot, B. B., 1982, *The Fractal Geometry of Nature* (San Francisco, California: Freeman).
[7] Meijas, P. M., and Martinez Herrero, R., 1991, *J. opt. Soc. Am.* A, **8,** 266–269.
[8] Berry, M. V., and Goldberg, J., 1988, *Nonlinearity*, **1,** 1–26.
[9] Averbukh, I. Sh., and Perelman, N. F., 1989, *Phys. Lett.* A, **139,** 449–453.

[10] Mallalieu, M., and Stroud, C. R., Jr., 1995, *Phys. Rev.* A, **51,** 1827–1835.
[11] Leichtle, C., Averbukj, I. Sh., and Schleich, W. P., 1996, Preprint, University of Ulm.
[12] Neponen, E., and Turunen, J., 1993, *Optics Commun.*, **98,** 132–140.
[13] Abramowitz, M., and Stegun, I. A., 1972, *Handbook of Mathematical Functions* (Washington, DC: US National Bureau of Standards).
[14] Patorski, K., 1983, *Optica Acta*, **30,** 1255–1258.
[15] Latimer, P., and Crouse, R. F., 1992, *J. opt. Soc. Am.*, **31,** 80–89.
[16] Smirnov, A. P., 1977, *Opt. Spectrosc. (USSR)*, **43,** 446–448.
[17] Lang, S., 1970, *Algebraic Number Theory* (Reading, Massachusetts: Addison-Wesley).
[18] Hannay, J. H., and Berry, M. V., 1980, *Physica* D, **1,** 267–290.
[19] Wolfram, S., 1991, *Mathematica: a System for Doing Mathematics by Computer* (Reading, Massachusetts: Addison-Wesley).
[20] Wronkowski, L., 1988, *J. mod. Optics*, **35,** 1417–1422.
[21] Arrizon, V., and Lopez-Olagasti, E., 1995, *J. opt. Soc. Am.* A, **12,** 801–804.
[22] Smirnov, A. P., 1978, *Opt. Spectrosc. (USSR)*, **44,** 208–212.
[23] Berry, M. V., and Lewis, Z. V., 1980, *Proc. R. Soc.* A, **370,** 459–484.
[24] Hardy, G. H., and Wright, E. M., 1988, *An Introduction to the Theory of Numbers*, fifth edition (Oxford: Clarendon).
[25] Berry, M. V., and Upstill, C., 1980, *Prog. Optics*, **18,** 257–346.
[26] Born, M., and Wolf, E., 1959, *Principles of Optics* (Oxford: Pergamon).
[27] Smirnov, A. P., 1979, *Opt. Spectrosc. (USSR)*, **46,** 319–322.
[28] Clauser, J. F., and Reinsch, M. W., 1992, *Appl. Phys.* B, **54,** 380–395.
[29] Falconer, K., 1990, *Fractal Geometry: Mathematical Foundations and Applications* (New York: Wiley).
[30] Berry, M. V., 1996, 'Quantum fractals in boxes', submitted to *J. Phys A*.