Aberration-like cusped focusing in the post-paraxial Talbot effect

James D Ring1, Jari Lindberg1, Christopher J Howls2 and Mark R Dennis1

1 H H Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, UK
2 School of Mathematics, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

E-mail: james.ring@bristol.ac.uk

Received 3 April 2012, accepted for publication 7 June 2012
Published 21 June 2012
Online at stacks.iop.org/JOpt/14/075702

Abstract

We present an analysis of the Talbot effect beyond the paraxial regime, where deviation from Fresnel propagation destroys perfect, periodic self-imaging. The resulting interference structures are examples of aberration without geometric optical rays, which we describe analytically using post-paraxial theory. They are similar to, but do not precisely replicate, a standard integral representation of a diffraction cusp (the Pearcey function). Beyond the Talbot effect, this result illustrates that aberration—as the replacement of a perfect focus with a cusp-like pattern—can occur as a consequence of improving the paraxial approximation, rather than due to imperfections in the optical system.

Keywords: post-paraxial, Pearcey function, surface plasmons

The wave-optical effect of aberration on foci is described in terms of diffraction catastrophes [18], which tend to ray caustics in the geometrical optics limit. In the case of two-dimensional (2D) ‘spherical aberration’ (i.e. cylindrical aberration), the diffraction pattern is described by the Pearcey function [19, 20], which is the interference pattern decorating a ray cusp. This pattern is particularly clear for spherical microlenses [21], illustrating it as the universal pattern accompanying spherical aberration, on the wavelength scale.

A similar interference pattern occurs instead of Talbot revivals when the transverse period is approximately 20 wavelengths, as evident in figure 1. However, the usual Talbot effect is lensless imaging and the distortion of the foci rather occurs as terms beyond Fresnel effectively aberrate the perfect paraxial foci. Furthermore, a geometric optical ray approximation is not appropriate for the initial field we consider, since it is an array of 2D point sources and not a phase varying field with uniform intensity [22]. In the regime where the period is close to the wavelength, aberration is not ‘spherical’ (i.e. from imperfect lenses), but rather originates from an improvement of the approximation.

Mathematically, a Pearcey function generically appears in place of a diffractive focus whenever a quadratic phase...
variation in the initial field is accompanied by a small quartic variation. This is independent of the physical origin of the phase variation, whether from the imperfection of a lens, as in usual aberration, or here in a diffractive regime where the Fresnel approximation is not appropriate.

In this paper, we invoke the fourth-order post-paraxial approximation [12, 14, 15] to investigate the Talbot-based focusing of scalar waves in a regime which is nonparaxial, but not extremely so: the transverse periodicity is of the order of some tens of wavelengths. This explains the appearance of diffraction cusps in the intensity patterns of surface plasmons calculated in [5]. We show that these aberration-like effects are not unique to plasmons and are connected to the standard integral representation of the diffraction pattern about a cusped caustic (the Pearsley function [18]) despite the absence here of any actual ray caustic. Furthermore, we obtain a simple expression for the approximate location of the post-paraxial foci, which are shifted to be short of the paraxial Talbot distance.

It might seem strange to some readers that we are making an approximation to a field whose analytical form is already known, and is simply a finite sum of plane waves. Nevertheless, the interference structures which are present in the pattern are not directly evident from the sum, but are emergent after the mathematical procedure we describe. The approximation procedure itself provides insight into the underlying physics.

To demonstrate the physics of the post-paraxial regime, we employ a general 2D formalism which displays properties of both paraxial and nonparaxial regimes in different limits. Specifically, it consists of a transverse array (with period $a$ in the $x$-direction) of monochromatic 2D point sources of wavelength $\lambda$. The interference of the propagating waves in the $z$-direction weaves a diffraction carpet in the $x$-$z$ plane, which is similar to that studied in the plasmon Talbot effect of [5]. The paraxial approximation applies when $\gamma = a/\lambda$ is asymptotically large. In this case the Talbot distance $2a^2/\lambda$ is the spacing between planes of image revival. The initial pattern also revives at odd multiples of $a^2/\lambda$, albeit transversely shifted by $a/2$, leading to an alternate definition of the Talbot distance [15] of $\gamma T \equiv a^2/\lambda$ (used in this paper). In what follows, we use the dimensionless variables $X = x/a$ and $Z = z/\gamma T$. In general, for any $\gamma$, the field at a distance $Z$, propagating according to the 2D Helmholtz equation, is given by (cf equation (1) of [5])

$$\psi(X, Z) = \sum_{n=-N}^{N} \exp \left[ 2\pi i \left( nX + \gamma^2 Z \sqrt{1 - n^2/\gamma^2} \right) \right],$$  

(1)

where we assume negligible attenuation and consider only propagating waves, i.e., $N = \lfloor \gamma \rfloor$ with square brackets denoting the integer part. Equation (1) can be approximated by expanding

$$\sqrt{1 - n^2/\gamma^2} \approx 1 - n^2/2\gamma^2 - n^4/8\gamma^4.$$

(2)

In the paraxial approximation to equation (1), the quartic term in equation (2) is negligibly small, so

$$\psi_p(X, Z) = \sum_{n=-N}^{N} \exp[i(2\pi nX - \pi n^2 Z)],$$

(3)

where we have neglected the overall phase factor of $\exp(2\pi i z/\lambda)$. As shown in figure 1(a), the initial paraxial field revives for integer $Z$, since the second term in equation (3) will introduce a factor of $\pm 1$, and we recover the same field as for $Z = 0$. The factor of $\pm 1$ determines whether or not the self-image is shifted by half a period. A
post-paraxial refinement is made by retaining the quartic term in equation (2) [12, 14, 15], giving

\[
\psi_{pp}(X, Z) = \sum_{n=-N}^{N} \exp\left[i(2\pi n X - \pi n^2 Z - \pi n^4 Z/4\sqrt{2})\right].
\]

Comparison between equations (4) and (3) shows that the perfect Z-periodicity of the Talbot effect is an artefact of the paraxial approximation, and is destroyed in the post-paraxial regime by the inclusion of the quartic (and possibly higher terms) [15]. Figures 1(b) and (d) show the exact and post-paraxial intensity distributions respectively, and it is immediately apparent that while the foci of these fields are similar to one another, they are very different, and less defined, from those of the paraxial case, (a). Comparing the foci of each regime in figure 1 it is clear that the perfect foci of the paraxial approximation of (a) have been perturbed in the post-paraxial regime to cusp-like distributions, evident in (b) and (d). Figure 1(c) shows a separate simulation for a binary grating, which has been calculated numerically using rigorous electromagnetic theory [17, 23] (parameters given in the caption), which very closely resembles the exact and post-paraxial fields in (b) and (d), including the cusp-like foci. While such a pattern is typical for aberration, its appearance here is because of a refinement of the usual Fresnel paraxial approximation. In fact, the inclusion of a quartic term destroying a perfect focus is exactly analogous to the spherical aberration of a parabolic mirror, though remarkably, in figure 1 it is improvement of the approximation and not any aberration in the system that causes the effect!

The fields in figure 1 are those for an infinite array, and applying the paraxial approximation to such a system is not fully consistent [16], since at all Z there would be oblique contributions to the field from arbitrarily transversely distant sources—a fact that is at odds with the concept of paraxiality. Instead we model a finite array of N sources by truncating the sums of equations (1), (3) and (4) at \pm M, where \(M \equiv [N/2\pi]^{[2]} [15, 24]\, embodying the walk-off effects of finite arrays [2]. Figure 2 shows the various foci according to this truncation. Comparing the paraxial focus figure 2(a) to the exact figure 2(b), the aberration effect is clear. Figure 2(c) shows the cusp pattern from the diffraction grating of figure 1(c), but includes only diffraction orders up to \pm M. There is a slight discrepancy between (b) and (c) because of the equal amplitudes of the planes waves in the superposition of equation (1) compared to varying amplitudes of the diffraction orders of the simulation of (c). The fields are nevertheless very similar and the cusp reveals the breakdown of the paraxial approximation. Below, we develop an analytic post-paraxial approximation to these fields, shown in figures 1(d), 2(d), from inclusion of the quartic term (and no higher), which is enough to capture the structure of the exact and simulated foci. Including higher terms improves the quantitative agreement, but without further qualitative changes.

To derive a simple representation of the complicated intereference pattern, we follow a similar technique to [15] but employ a finite Poisson sum formula [25] to account for the truncation at \pm M of the spectrum. Applying this to \(\psi_{pp}(X, Z)\) yields the main result of this paper,

\[
\psi_{pp}(\xi, \zeta) = \frac{f(-M; \xi, \zeta)}{2} + \frac{f(M; \xi, \zeta)}{2} + \sum_{\mu=-\infty}^{\infty} \int_{-\tau(\xi)}^{\tau(\xi)} ds \, g(s; \xi, \zeta, \mu),
\]

where \(g(s; \xi, \zeta, \mu) \equiv h(\zeta) \exp[-i(s^4 + \sigma_{s}(\xi, \zeta)s^2 + \sigma_{s}(\xi, \mu, \zeta)s)],\) and

\[
f(n; \xi, \zeta) \equiv \exp[\pi\sigma(2\xi n - \zeta n^2 - (T + \zeta) n^2/4\sqrt{2})];
\]

\[
h(\zeta) \equiv \sqrt{2\sqrt{y}/[\pi(T + \zeta)]^2};
\]

\[
\tau(\zeta) \equiv M/h(\zeta);
\]

\[
\sigma_{s}(\xi, \mu, \zeta) \equiv 2\pi(\xi - \mu) h(\zeta);
\]

\[
\sigma_{s}(\zeta) \equiv \pi h^2(\zeta) \zeta.
\]

In the above, we have set \(Z = T + \zeta,\) where \(\zeta\) is a small shift from the 7th Talbot distance, and \(\xi = X - T/2.\) The terms \(f(\pm M; \xi, \zeta)\) are end-point contributions from the truncation of the spectrum. We recognise that the integral in equation (5) is an incomplete version [21] of the Pearcey function [19, 20]
(i.e. with a finite integration domain). The Pearcey function is given by

\[ \text{Pe}(x, z) = \int_{-\infty}^{\infty} ds \exp[i(s^4 + zs^2 + xs)], \]

which describes the generic diffraction pattern decorating a cusp caustic [18, 20].

The sum in equation (5) represents an array of cusp-like foci in the x-direction, with transverse and longitudinal scalings given by \( \sigma_{x} \) and \( \sigma_{z} \). Although in optics it is usual to see the Pearcey function as decoration around a cusp caustic of rays that arise from a geometrical optics approach [18], this is not necessary [22]. Since here we are outside the applicability of geometrical optics, the propagating function equation (1) cannot be approximated in terms of rays, and instead the Pearcey function has come from the quartic post-paraxial term in equation (4). Importantly, since the maximum of the Pearcey function is below the geometric focus, the maxima of the post-paraxial self-images are below the paraxial Talbot distance. This differs from the nonparaxial Talbot distance [8, 17] which is accurate when only a small number of diffraction orders contribute.

Figure 2(d) shows the diffraction pattern near the Talbot distance (\( Z = 1 \)) according to equation (5) for \( |\mu| \leq 1 \). The intensity of the Pearcey function, scaled to the size set by \( \sigma_{x} \) and \( \sigma_{z} \), is shown in figure 2(e). Figure 2(f) is the incomplete Pearcey function (the \( \mu = 0 \) term of equation (5)). Comparing (b) and (d), we see that although the sum in equation (5) has infinitely many terms, only the first few are needed to accurately describe the fine details of the field about the focus. The transverse fringes come from the truncation of the integral in equation (5), and subsequently are absent from the Pearcey function, (e), and are just evident in (f). In all these images, it is clear that the maximum intensity is short of the paraxial Talbot distance. We now obtain an expression for the approximate location of the maximum intensity.

As a first approximation we take only the \( \mu = 0 \) term of equation (5) and let \( \tau(\xi) \) go to infinity. This is equivalent to approximating the focal region near the first Talbot distance with the Pearcey function, equation (7). By using the numerically calculated, constant position of maximum intensity of the unscaled Pearcey function, \( Z_{\text{Pe}} \approx -2.199 \), we can employ the scalings of equation (6) to calculate the approximate location of the maximum intensity of the post-paraxial focus as a shift, \( \zeta_{\text{peak}} \), from the \( T \)th Talbot distance. By setting \( \sigma_{z}(\zeta_{\text{peak}}) = Z_{\text{Pe}} \), we obtain

\[ \zeta_{\text{peak}} = \frac{Z_{\text{Ps}}^2 + Z_{\text{Pe}}^2 \sqrt{Z_{\text{Ps}}^2 + 16\pi^2 \gamma^2 T}}{8\pi \gamma^2}. \]

Figure 3(a) shows the intensity of the exact field along the symmetry axis of the cusp \( \xi = 0 \) for \( -0.1 \leq \zeta \leq 0.05 \) as a function of increasing \( \gamma \). Asymptotically large \( \gamma \) gives rise to maximum intensity at the Talbot distance (\( \zeta = 0 \) in the paraxial limit), but for smaller \( \gamma \) it is located short of this. Figure 3(b) shows the post-paraxial approximation of equation (5) including only the terms \( |\mu| \leq 1 \). The dashed curve is the location of the maximum intensity according to the approximation of equation (8). Note that this depends on \( T \), such that consecutive foci will be shifted from their paraxial counterparts by an increasing amount. The dots indicate the maxima of the exact field (equation (1)) for integer \( \gamma \), though they do not always correspond to well localised single peaks, as in the case of \( \gamma = 12 \) and 28. The intensity fringes that we observe in figure 3 are a result of interference between the \( \mu = 0 \) and the neighbouring \( \mu = \pm 1 \) terms. Inclusion of more \( \mu \)-terms reconstructs the discontinuities of figure 3(a) (which arise from the limits of the sum in equation (1) changing discretely with \( \gamma \)).

To conclude, we have applied scalar post-paraxial theory to explain the observed cusp-like foci in the Talbot effect for the post-paraxial regime, which have replaced the point foci of paraxial theory. While the analytic form of these post-paraxial foci is related to the standard description of diffraction about a cusp, there is a subtle distinction since it is the incomplete Pearcy function that occurs and not the standard Pearcey function. The finite integration domain of the incomplete Pearcy is of significant asymptotic importance when compared with the infinite domain of the complete Pearcy function of equation (7). This is clear when comparing intensity patterns—while both are cusp-like distributions, there is additional complex detail accounted for in the post-paraxial Talbot effect that is absent from the normal Pearcey function. Another important difference is that there is no geometric optical ray caustic underpinning the structure of the post-paraxial cusp foci as there is with the complete Pearcey function. Despite their differences, the location of the maximum of intensity of the post-paraxial foci can be fairly well approximated using a scaled Pearcey function.

Up to rescaling, the Pearcey function universally describes cylindrically aberrated 2D foci. Although Pearcey function revivals have been studied previously in the Talbot effect [22], they occurred there in paraxial propagation as the result of specifically chosen initial conditions, rather than as revived foci whose aberration follows from propagation in the post-paraxial regime.
Furthermore, aberrations following from nonparaxial corrections have also been studied with an emphasis on the geometrical optics limit [12, 13], where it was previously found that post-paraxial corrections cause breakdown of perfect self-imaging. Here we emphasise the analytic approximation of the aberrated diffraction pattern, described by the incomplete Pearcey function.

It is curious to see that the effect of post-paraxiality on a focus appears to be that of aberration, when compared to the paraxial case. However, while the analytic form mimics that for aberration (hence the inclusion of the Pearcey function, albeit incomplete), it actually embodies an improvement of the paraxial approximation for the post-paraxial regime. The phenomenon is a universal effect that occurs as the parameters of a system (regardless of its exact realisation) move away from paraxiality.

The results in this paper may plausibly find application in plasmonics, since it is a field in which the pertinent wavelengths and periodicities are comparable, and so nonparaxial effects are important. While the lengthscales considered here are slightly greater than those experimentally investigated so far, it would be both interesting and challenging to confirm the aberration-like effects of post-paraxiality on the Talbot effect.

Acknowledgments

We gratefully acknowledge the funding of the EPSRC. MRD is a Royal Society University Research Fellow.

References

[1] Talbot H F 1836 Facts relating to optical science. No. IV Phil. Mag. 9 401–7
[2] Patorski K 1989 The self-imaging phenomenon and its applications Prog. Opt. 27 1–108
[3] Chapman M, Ekstrom C, Hammond T, Schmiedmayer J, Tannian B, Wehinger S and Pritchard D 1995 Near-field imaging of atom diffraction gratings: the atomic Talbot effect Phys. Rev. A 51 R14–17
[4] Averbukh I and Perelman N F 1989 Fractional revivals: universality in the long-term evolution of quantum wave packets beyond the correspondence principle dynamics Phys. Lett. A 139 449–53
[5] Dennis M R, Zheludev N and García de Abajo F 2007 The plasmon Talbot effect Opt. Express 15 9692–700
[6] Martinez Niconoff G, Sanchez-Gil J A, Sanchez H H and Perez Leija A 2008 Self-imaging and caustics in two-dimensional surface plasmon optics Opt. Commun. 281 2316–20
[7] Maradudin A A and Leskova T A 2009 The Talbot effect for a surface plasmon polariton New J. Phys. 11 033004
[8] Zhang W, Zhao C, Wang J and Zhang J 2009 An experimental study of the plasmonic Talbot effect Opt. Express 17 19757–62
[9] Cherukulappurath S, Heimis D, Cesario J, van Hulst N, Enoch S and Quidant R 2009 Local observation of plasmon focusing in Talbot carpets Opt. Express 17 23772–84
[10] Wan X, Wang Q and Tao H 2010 Nanolithography in the quasi-far field based on the destructive interference effect of surface plasmon polaritons J. Opt. Soc. Am. A 27 973–6
[11] Liu Z, Wei Q and Zhang X 2005 Surface plasmon interference nanolithography Nano Lett. 5 957–61
[12] Cohen-Sabbah Y and Joyeux D 1983 Aberration-free nonparaxial self-imaging J. Opt. Soc. Am. 73 707–19
[13] Chang S 2005 Geometrical aberrations of self-imaged line gratings Optik 116 379–89
[14] Leger J and Swanson G 1990 Efficient array illuminator using binary-optics phase plates at fractional-Talbot planes Opt. Lett. 15 288–90
[15] Berry M V and Klein S 1996 Integer, fractional and fractal Talbot effects J. Mod. Opt. 43 3139–64
[16] Mejias P M and Martinez Herrero R 1991 Diffraction by one-dimensional Ronchi grids: on the validity of the Talbot effect J. Opt. Soc. Am. A 8 266–9
[17] Noponen E and Turunen J 1993 Electromagnetic theory of Talbot imaging Opt. Commun. 98 132–40
[18] Nye J F 1999 Natural Focusing and Fine Structure of Light (Bristol: Institute of Physics Publishing)
[19] Pearcey T 1946 The structure of an electromagnetic field in the neighbourhood of a cusp of a caustic Phil. Mag. Series 7 37 311–7
[20] Berry M V and Howls C J Digital Library of Mathematical Functions: Catastrophes and Canonical Integrals http://dlmf.nist.gov/36.2
[21] Nye J F 2003 Evolution from a Fraunhofer to a Pearcey diffraction pattern J. Opt. A 5 495–502
[22] Berry M V and Bodenschatz E 1999 Caustics, multiply reconstructed by Talbot interference J. Mod. Opt. 46 349–65
[23] Moharam M G, Grann E B, Pommet D A and Gaylord T K 1995 Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings J. Opt. Soc. Am. A 12 1068–76
[24] Berry M V and Goldberg J 1988 Renormalisation of curlicues Nonlinearity 1 1–26
[25] Civi Ö, Pathak P and Chou H 1999 On the Poisson sum formula for the analysis of wave radiation and scattering from large finite arrays IEEE Trans. Antennas Propag. 47 958–9