Review article

Greg Gbur*

Using superoscillations for superresolved imaging and subwavelength focusing

https://doi.org/10.1515/nanoph-2018-0112
Received August 3, 2018; revised October 24, 2018; accepted October 30, 2018

Abstract: It is now well-appreciated that a bandlimited wave can possess oscillations much more rapidly than those predicted by the bandlimit itself, in a phenomenon known as superoscillation. Such superoscillations are required to be of dramatically smaller amplitude than the signal they are embedded in, and this has initially led researchers to consider them of limited use in applications. However, this view has changed in recent years and superoscillations have been employed in a number of systems to beat the limits of conventional diffraction theory. In this review, we discuss the current state of research on superoscillations in terms of superresolved imaging and subwavelength focusing, including the use of special nondiffracting and Airy beams to carry transverse superoscillating patterns. In addition, we discuss recent analogous works on using superoscillations to break the temporal resolution limit, and also consider the recently introduced inverse of superoscillations, known as suboscillations.

Keywords: superoscillations; singular optics; superresolution; imaging.

1 Introduction

Traditionally, it has long been assumed that a bandlimited signal has a maximum oscillation frequency, in space or in time, naturally dictated by the largest frequency in the band. For example, if a temporal signal \( f(t) \) is bandlimited, such that its Fourier spectrum,

\[
\hat{f}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \, dt,
\]

is identically zero outside the range \(-\Delta \leq \omega \leq \Delta\), then conventional wisdom suggests that the signal itself will have no oscillations with a period shorter than \(2\pi/\Delta\). However, it is now well recognized that this is not a strict limit, and it is possible for a signal to possess superoscillations, i.e., regions in which the local frequency is larger, in principle arbitrarily larger, than the highest nonzero frequency in the band.

An understanding of such superoscillations comes from the recognition that the zeros of a function may be packed arbitrarily close together, regardless of the bandwidth of the function itself. The space between a pair of closely-spaced zeros represents a half-oscillation of a signal, and therefore the half-period of a function may be made arbitrarily small. Furthermore, as many zeros as desired may be placed near each other, in principle resulting in as many superoscillations as desired in a given region.

These superoscillations come with a cost, however. The amplitude between a pair of closely-spaced zeros is small compared to that of the surrounding function; in fact, it gets even smaller the closer the zeros are packed. In an early introduction to the concept of superoscillations, Berry [1] estimated that reproducing Beethoven’s Ninth Symphony using superoscillations in a 1 Hz bandlimited signal would require a background signal amplified by a factor of \(10^{19}\), which is clearly impractical as a signal-compression technique.

A short history of the subject is worth summarizing, as many insights come from some of the earliest works. The first results originated in the theory of linear antenna
arrays in the 1940s, as studied by Schelkunoff [3]. A linear array is one in which a finite number of identical, equally-spaced antennas are arranged along a straight line; the only difference between elements of the array are their relative amplitudes and phases. In laying out an elegant mathematical theory, Schelkunoff noted that it is possible to make arrays that have, apparently, arbitrarily high directionality. This is done by packing zeros of the radiation pattern (known as the “cones of silence”), densely around the directional peak. In essence, the zeros are used to push the sidelobes of the radiation pattern out of the regime of the propagating waves. This work received much attention and led to the search for the highest directionality that can be achieved for such a strategy; in 1946, Bouwkamp and de Bruijn demonstrated [4] that there is no “optimal” solution to the problem, implying that there is no upper limit to directionality, at least within the constraints of the problem as stated.

New constraints were soon found. In 1948, Woodward and Lawson [5] pointed out that the amount of current required in the individual elements becomes impractical for all but the most modest improvements in directivity; this was later confirmed by Yaru [6]. In essence, the sidelobes that are pushed out of the propagating region of the antenna pattern become massive and represent a massive reactive current contribution. Nevertheless, in 1952, Toraldo di Francia drew an analogy between these so-called super-gain antennas and the resolving power of optical systems [7], hinting at the possibility of superresolution in optics.

The existence of rapid oscillations of a bandlimited signal was next noted in the context of signal processing. For example, in 1977, Khurgin and Yakovlev argued [8] that

We thus arrive at the conclusion that one can transmit a signal of arbitrarily high frequencies by means of a low-frequency signal. To do this, however, the total energy of the transmitted low frequency signal must grow.

This description points out the advantages and disadvantages of superoscillations already discussed. In 1986, Landau [9] looked at the possibility of extrapolating a band-limited function from a collection of oversamples. He noted that, due to the possibility of close-packed zeros, it is necessary to add additional constraints to make the problem a well-posed one. In essence, any set of oversampling points may coincide with the zeros of a superoscillatory function; obviously, that function would not provide any contribution to the extrapolation. To overcome this issue, Landau required that the energy of the total signal be bounded – a strategy that would restrict the size of any superoscillating components of a solution.

As a wave phenomenon of purely physical interest, superoscillations were first considered by Aharonov et al. in the introduction of weak values of a quantum variable [10]. In a follow-up study, they theoretically demonstrated [11] a method to obtain a superposition of different Hamiltonian time evolution operators, and showed that this method can act effectively like a single evolution operator whose behavior was completely out of the range of the original set of operators. This is a quantum operator analogy of adding together a collection of time-harmonic signals to get a signal that oscillates faster than any of its components. Aharonov pointed out this work to Michael Berry, who wrote a paper dedicated to Aharonov [1], in which he popularized the term “superoscillations,” and this paper led to broader recognition of the phenomenon.

For the next decade, there was relatively little discussion of superoscillations in the literature. Among several papers of note is the 2000 work by Kempf [12], in which he considers further the possibility of data compression with superoscillations – Beethoven at 1 Hz – and also demonstrates an uncertainty relation for bandlimited functions that accounts for superoscillations. Further investigations of the basic mathematics of superoscillations were presented by Calder and Kempf [13] and Ferreira [14].

In 2006, Berry and Popescu [15] presented the first discussion of superresolution via superoscillations, in which they demonstrate that superoscillations are surprisingly robust on propagation and therefore can carry sub-wavelength structure beyond the range of evanescent waves. Spurred by this insight, multiple researchers began in earnest to investigate the possibility of using superoscillations in applications, and this is the subject of the current review.

In this article, we will discuss the progress made in the application of superoscillations to nano-optics imaging applications. In Section 2, we introduce the basic mathematics of superoscillations, providing a variety of different examples and methods of constructing such oscillations. In Section 3, we review the use of superoscillations in the production of subwavelength focal spots. In Section 4, we look at the progress made in using superoscillations in superresolved imaging. In Section 5, we review the structured beams that have been designed to carry superoscillations without diffraction over long distances. In Section 6, we look at notable advances in superoscillations, including the development of complementary “suboscillations” that oscillate slower than the lowest frequency in the band. Finally, in Section 7, we present some concluding remarks.
2 Mathematics of superoscillations

There are many mathematical techniques for illustrating and designing superoscillations in waves. In this section, we look at a number of these techniques, as different models have proven to be useful in different circumstances.

We begin with the precursor to superoscillations, namely, super-gain in antennas; more information can be found in Berry [16]. The geometry of a linear antenna array is illustrated in Figure 1. There are $N$ identical antennas arranged along a straight line; the first antenna is labeled 0 and the last antenna is therefore labeled $N-1$. The radiation pattern of the antenna includes the response of an individual antenna as well as the interference effects between them. In the far zone, the individual response factors out of the expression for the field and we may focus on the interference pattern, called the “space factor”. In the original work, each antenna is assumed to have a progressive phase delay of $-\Delta \phi$ relative to its left neighbor. The shape of the radiation pattern can be tuned by adjusting the relative amplitude and phase of its left neighbor. The result is observed in the far zone at an angle $\theta$ relative to the end-fire antenna.

In the far zone, the radiating elements may focus on the interference pattern, called the “space factor” in the original work. Each antenna is assumed to have a progressive phase delay of $-\Delta \phi$ relative to its left neighbor. The shape of the radiation pattern can be tuned by adjusting the relative amplitude and phase of its constituent elements, represented by a complex number $a_n$. The total space factor $U(\theta)$ for the entire array may be written as

$$U(\theta) = \sum_{n=0}^{N-1} a_n e^{j[kl\cos \theta - \Delta \phi]} = \sum_{n=0}^{N-1} a_n z^n,$$  \hspace{1cm} (2)

where we used $z = \exp(j[kl\cos \theta - \Delta \phi])$.

As $0 \leq \theta \leq \pi$, the argument of $z$ ranges from $-k\lambda - \Delta \phi \leq \arg(z) \leq k\lambda - \Delta \phi$. Given that $|z|=1$, the values of $z$ represent an arc on the unit circle. If the elements are spaced by a half-wavelength, i.e. $l = \lambda/2$, then $z$ covers the entire unit circle; for smaller spacings, the range is smaller.

This formulation, by Schelkunoff [3], is a remarkably elegant way to characterize any linear array in terms of analytic polynomials whose properties are well-understood. By the fundamental theorem of algebra, an array of $N$ elements may always have its space factor written, within a multiplicative constant, in the form

$$U(z) = (z-z_1) \cdots (z-z_{N-1}),$$  \hspace{1cm} (3)

where $z_1, \ldots, z_{N-1}$ are the zeros of the polynomial. If those zeros lie on the unit circle, they potentially represent zeros of the radiation pattern, but only if they lie in the range of the argument of $z$.

By choosing the position of these zeros appropriately, one can significantly improve the directionality of an array. Let us restrict ourselves now to the case $l = \lambda/4$; the range of $z$-values, therefore, covers half of the unit circle. We further take $\Delta \phi = kl$, which places $\theta = 0$ at $z = 1$, and makes the range of $z$ the lower half of the unit circle; this is the case of an end-fire antenna.

We now consider the radiation patterns of two antennas. The simplest example is a uniform array, for which $a_n = 1$. The space factor is just a finite geometric series, and may be written as

$$U(z) = \sum_{n=0}^{N-1} z^n = \frac{z^N - 1}{z - 1}.$$  \hspace{1cm} (4)

For this array, the zeros are the $n$-th roots of $z$, and are equispaced around the unit circle. In terms of intensity, the space factor may be written as

$$|U(z)|^2 = \left[ \frac{\sin(N\psi/2)}{\sin(\psi/2)} \right]^2,$$  \hspace{1cm} (5)

where $\psi = (\cos \theta - 1)/2$ is the argument of complex $z$. The peak value will be at $\theta = 0$, and the first zero of the radiation pattern will appear when $z = \exp(-2\pi i/N)$, or $\cos \theta = 1 - 4/N$.

However, as Schelkunoff noted, we can improve the directionality of the array by compressing all of the zeros of the function into the lower half of the unit circle, with a space factor

$$U(z) = (z-e^{-\pi i/N})(z-e^{-2\pi i/N}) \cdots (z-e^{-(N-1)\pi i/N}).$$  \hspace{1cm} (6)

An illustration of the radiation patterns of a uniform array and the compressed array are shown in Figure 2. One can see that the array with compressed zeros has a narrower primary peak as well as smaller sidelobes, indicating a significant overall improvement in directionality.

An illustration of the cost of this improvement is shown in Figure 3. Plotting the space factor for the entire argument of $z$, which includes both the propagating components and the evanescent components of the field, one sees that there is a huge peak for the compressed array in the evanescent region. In order for the compressed array to

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{The notation related to a linear antenna array.}
\end{figure}
G. Gbur: Superoscillations for imaging

have the same peak radiated power as the uniform array, one must put more total power into the compressed array overall. Essentially, this same issue appears for superoscillations in optical waves at finite propagation distances.

The first illustrative model of superoscillations in optics was presented by Berry [1], based on an argument by Aharonov. We consider the function

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2\delta^2} \left( u - iu_c \right)^2 \right] du,$$

where

$$k(u) = \frac{1}{1 + u^2},$$

and \( \delta \) and \( u_c \) are real-valued parameters. Here, the exponential \( \exp[ik(u)x] \) looks roughly like a Fourier kernel with a maximum spatial frequency of unity; by an appropriate coordinate transformation, the entire integral can be rewritten as a function bandlimited with \(|k| < 1\). The second exponential is of Gaussian form, with the center of the Gaussian in the complex plane at position \( iu_c \). As \( \delta \to 0 \), this function will act like a Dirac delta function, centered on the imaginary axis. The argument proposed by Aharonov is that this delta will nevertheless effectively “sift” out this complex part of the integral, resulting in a function locally oscillating with a spatial frequency expressed as

$$k(iu_c) = \frac{1}{1 - u_c^2}.$$  \( \tag{9} \)

Clearly, with \( 0 < u_c < 1 \), this results in oscillations that are faster than those allowed by the bandlimit (i.e. superoscillations). In fact, we may see superoscillations for any \(|u_c| < 2\), with \(|u_c| \neq 1\). Berry demonstrated through asymptotic analysis that these superoscillations do, in fact, arise for \( g(x) \) for sufficiently small \( \delta \).

A much simpler model of a superoscillatory function was presented by Qiao [17]. It follows from the Paley-Wiener theorem that a function bandlimited in \( k \)-space is entirely analytic in \( x \)-space; from this, one can readily prove that changing the \( x \) position of a zero by a finite amount does not affect either the analytic properties of the function or its bandwidth. Then, any bandlimited function with known zeros may be made into a superoscillatory function by moving those zero positions. For example, the sinc function

$$f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikx} dk = \frac{\sin(\pi x)}{\pi x} = \prod_{n=1}^{\infty} \left( 1 - \frac{x^2}{n^2} \right)$$

is a bandlimited function with constant spectrum in the range \(-\pi \leq k \leq \pi\). Given that the zeros are well-defined by the rightmost part of this expression, a finite number \( N \) of them may easily be shifted closer together by a factor of \( \alpha > 1 \) to make a superoscillatory function \( g(x) \) of the form

$$g(x) = \prod_{n=1}^{N} \left( 1 - \frac{\alpha^2 x^2}{n^2} \right) \prod_{m=N+1}^{\infty} \left( 1 - \frac{x^2}{m^2} \right).$$

The superoscillations will appear in the regime, where \(-N/\alpha \leq x \leq N/\alpha\), as can be directly seen from the placement of zeros.

A simple example of this behavior, with \( N=2 \) and \( \alpha = 4 \), is shown in Figure 4. The natural period of the sinc function is \( \Delta x = 2 \); however, that period has quadrupled within the region of superoscillations. However, as can be seen from the right side of the figure, which shows the
complete vertical scale, the function \( g(x) \) has extremely large sidelobes flanking the region of superoscillation.

One can argue that this arises from the mismatch of the zero-shifted polynomials with the original polynomials. To see this, we note that

\[
g(x) = \prod_{n=1}^{N} \left( 1 - \frac{x^2}{n^2} \right) \frac{\sin(\pi x)}{\pi x}.
\]

Asymptotically in \( x \), the ratio of polynomials will tend to \( \alpha^N \), and the function will appear as the original sinc function scaled by this amount. However, for the intermediate values of \( x \), the numerator will blow up to extremely large values.

An extremely simple superoscillatory function was discussed in detail in 2006 by Berry and Popescu [15], though it was first introduced by Aharonov et al. [10]. Sometimes referred to as the canonical superoscillatory function, today, it is of the form

\[
g(x) = \left( \cos x + i a \sin x \right)^N,
\]

with \( N \gg 1 \) and \( a > 1 \). It is trivial to see that this function may be written in a Fourier series representation, where the fastest oscillations have a spatial frequency \( N \), due to the appearance of \( (\cos x)^N \) and \( (\sin x)^N \). If we consider its behavior near \( x = 0 \), however, we may approximate \( \cos x \) and \( \sin x \) by the lowest order terms of their Taylor series, and write

\[
g(x) = (1 + iax)^N = \exp[N \log(1 + iax)] = \exp[iNax],
\]

where we further use the Taylor series approximation of the logarithm function. This function, therefore, exhibits superoscillations, and it is a simple enough function to evaluate many of its properties analytically. To do so, it is convenient to rewrite the function in a different analytic form, as Berry and Popescu did. We first express this in terms of an amplitude and phase term

\[
g(x) = \left| \cos x + i a \sin x \right|^N \frac{(\cos x + i a \sin x)^N}{(\cos x + i a \sin x)^N} = \left| \cos x + i a \sin x \right|^N \exp[iN\psi].
\]

Then the argument of the phase term is expressed as

\[
\psi = \arg \left\{ \frac{\cos x + i a \sin x}{(\cos x + i a \sin x)} \right\} = \arctan \left( \frac{a \sin x}{\cos x} \right).
\]

We write this argument as an integral

\[
\psi = \int_0^x \frac{d}{dx'} \arctan \left( \frac{a \sin x'}{\cos x'} \right) dx' = \int_0^x \frac{a}{\cos^2 x' + a^2 \sin^2 x'} dx'.
\]

If we define \( k(x) \) to be a local wavenumber of the form,

\[
k(x) = \frac{a}{\cos^2 x + a^2 \sin^2 x},
\]

then it can be seen in a straightforward manner that we may write

\[
g(x) = \left( \frac{a}{k(x)} \right)^{N/2} \exp \left[ iN \int_0^x k(x') dx' \right].
\]

From the definition of \( k(x) \), the fastest oscillations (at \( x = 0 \)) have a local wavenumber \( k_{\text{max}} = aN \), or a zero spacing \( \Delta x_{\text{min}} = \pi/(aN) \). The slowest oscillations (at \( x = \pi/2 \)) have a local wavenumber \( k_{\text{min}} = N/a \), and a zero spacing \( \Delta x_{\text{max}} = \pi a/N \). The amplitude in the region of the fastest oscillations can be found to be approximately unity, and the amplitude in the region of the slowest oscillations is \( a^N \).

From Eq. (19), it is clear that we can define a local frequency operator \( \kappa(x) \), which is of the form

Fig. 4: The sinc function \( f(x) \) and the superoscillatory version \( g(x) \) with \( N = 2, \alpha = 4 \), plotted with two different vertical scales.
\[ \kappa(x) = \text{Im} \left( \frac{d}{dx} \log(f(x)) \right), \]  
\text{(20)}

and takes the function \( g(x) \) as its argument and provides a local value of the spatial frequency at the point \( x \). This operator can be used to find the local superoscillatory regions of a waveform, and it has become a standard tool in such investigations.

The function \( g(x) \) is illustrated in Figure 5 with \( a = 3, N = 20 \). It is plotted on a semilog scale in order to clearly see those positions at which the amplitude drops dramatically, indicating a zero-crossing. The predicted maximum amplitude for this case is \( a^c = 3.5 \times 10^6 \), which is in agreement with the figure. Notably, the function is periodic, and by looking at a larger region, as in Figure 5B, one can see that the superoscillations are mixed in with the natural period of the functions \( \sin x \) and \( \cos x \).

In hindsight, superoscillations have been hiding in plain sight for decades, in the form of prolate spheroidal wave functions (PSWFs), as pointed out by Ferreira [14]. The PSWFs were studied extensively in the 1960s by Landau, Pollack and Slepian [18, 19]; a modern discussion was given by Moore and Cada [20]. Let us consider functions in \( x, k \) space. For a given bandlimit \( \Delta k \), such that \( |k| \leq \Delta k \) and an arbitrary interval \( \Delta x \) in space, it can be shown that one can find a countably infinite set of functions \( \psi_n(c, x) \), \( n = 0, 1, 2, \ldots \), where \( c = \Delta k \Delta x \), such that

1. The \( \psi_n(c, x) \) are bandlimited, orthonormal and a complete set over the real line \(-\infty < x < \infty\), with
\[ \int_{-\infty}^{\infty} \psi_n(c, x) \psi_m(c, x) dx = \delta_{nm}, \]  
\text{(21)}

2. Over the interval \(-\Delta x \leq x \leq \Delta x\), the functions are orthogonal and form a complete set, with
\[ \int_{-\Delta x}^{\Delta x} \psi_n(c, x) \psi_m(c, x) dx = \delta_{nm} \lambda_n(c), \]  
\text{(22)}

where \( \lambda_n(c) \) is an eigenvalue of the sinc operator
\[ \int_{-\Delta x}^{\Delta x} \psi_n(c, x') \frac{\sin(k(x-x'))}{\pi(x-x')} dx' = \lambda_n(c) \psi_n(c, x). \]  
\text{(23)}

The number of zeros in the interval \(-\Delta x \leq x \leq \Delta x\) is equal to \( n \).

This last property indicates the existence of superoscillations. This is because the number of oscillations in a finite interval increases linearly with \( n \), inevitably resulting in superoscillations beyond some critical value of \( n \). The PSWFs are relatively obscure in optics, which is likely due to their complicated nature compared to other orthogonal expansions. They have been applied successfully in superoscillation-based imaging, as we will soon see.

Perhaps the most elegant method for demonstrating the feasibility of superoscillations, and prescribing them readily, was introduced by Chremmos and Fikioris [21] in 2015. Let us imagine that we have a function \( f(x) \), which is bandlimited to \( |k| \leq 2\pi \), i.e. \( \tilde{f}(k) = 0 \) outside the range \(-2\pi \leq k \leq 2\pi \). We introduce a new function \( g(x) \) by multiplying \( f(x) \) by a finite polynomial expressed as
\[ g(x) = \sum_{n=0}^{N} a_n x^n f(x), \]  
\text{(24)}

where the coefficients are arbitrary. According to elementary Fourier theory, the spectrum \( \tilde{g}(k) \) of the function \( g(x) \) is given by
\[ \tilde{g}(k) = \frac{1}{2\pi} \int \sum_{n=0}^{N} a_n k^n \tilde{f}(k) e^{ikx} dx = \frac{1}{2\pi} \sum_{n=0}^{N} a_n \left( \frac{d}{dk} \right)^n \tilde{f}(k), \]  
\text{(25)}

Provided that the first \( N \) \( \text{th} \) derivatives of \( \tilde{f}(k) \) are continuous, the function \( \tilde{g}(k) \) will be well-behaved and have exactly the same range as \( \tilde{f}(k) \). The polynomial is essentially arbitrary, and it is clear that we can choose the

![Figure 5](image-url)

Figure 5: The Berry-Popescu superoscillating function with \( a = 3, N = 20 \), for a range (A) \(-0.5 \leq x \leq 0.5\) and (B) \(-5.5 \leq x \leq 5.5\).
coefficients $a_n$ to place the zeros of the function at any position in $x$, even arbitrarily close together. Therefore, this method by Chremmos and Fikioris allows one not only to prove the existence of superoscillations, but to design functions with any desired superoscillatory pattern.

A very simple example is shown in Figure 6, with the polynomial

$$x(x^2 - a^2)(x^2 - 4a^2),$$

with $a = 1$, and the spectrum

$$f(k) = \begin{cases} \cos(\pi k / 2)^2, & |k| \leq 1, \\ 0, & |k| > 1. \end{cases}$$

The maximum spatial frequency that appears in the function is $k = 1$, which corresponds to a period of $2\pi$ and a zero spacing of $\pi$; however, the zeros of the polynomial, which represent superoscillations, are spaced by unity.

It is important to acknowledge that we have, by no means, exhausted the possible methods of creating superoscillations. Chojnacki and Kempf [22] introduced even more techniques in a recent article by including methods for making superoscillatory functions multiplicatively, which are either periodic or square integrable. The authors further introduced a strategy for designing quantum potentials possessing at least one superoscillatory eigenstate. In 2013, Katzav and Schwartz [23] introduced a method for optimizing the superoscillatory yield of a signal, given a fixed range and frequency for the superoscillations.

It should be noted that we have, so far, restricted our discussion to superoscillations in one-dimensional systems. However, superoscillations can also be demonstrated in higher-dimensional bandlimited waves, such as free-propagating paraxial beams. The typical zeros of scalar wavefields, which manifest as lines of zero intensity in a three-dimensional (3D) space, have a circulating or helical phase structure that has led to them being called optical vortices [24]. Under perturbation, an $N$th order vortex can break into $N$ 1st order vortices on propagation, and it is clear that those zeros are arbitrarily close together during the breakup. Optical vortices are generic, or typical features of scalar waves, and when there are many zeros present in a wave, superoscillations are inevitable. For example, it was demonstrated by Dennis, Hamilton and Courtial [25] that 1/5 of the cross-sectional area of a random speckle field may be considered superoscillatory, i.e. have a local wavenumber higher than the free-space wavenumber $k$. Similar results were shown by Berry and Dennis for isotropic random waves of an arbitrary number of dimensions [26].

As shown by Smith and Gbur [27], the method of Chremmos and Fikioris can be extended to design superoscillations in the transverse plane of a bandlimited optical field. Locally, a generic left-handed optical vortex line, of total order $N$ which is parallel to the axis of propagation, is expressed as $\eta = x + iy$ by introducing a polynomial of the form

$$h(\eta) = \sum_{n=0}^{N} a_n \eta^n,$$

and one can place vortices in a transverse bandlimited field $f(x, y)$ by the multiplication

$$g(x, y) = h(\eta) f(x, y).$$

In spatial frequency space, the spectrum $\tilde{g}(k_x, k_y)$ is given by the expression,

$$\tilde{g}(k_x, k_y) = \sum_{n=0}^{N} a_n \left[ \frac{\partial}{\partial k_x} + i \frac{\partial}{\partial k_y} \right]^n \tilde{f}(k_x, k_y).$$

We may generalize the function $h(\eta)$ of Eq. (28) to include the right-handed vortices of the form $\eta' = x - iy$ by multiplying it by its own polynomial in $\eta'$. We may further include vortices of any orientation by using polynomial terms of the form $\eta = x + iby$, with $b$ a real number.
Such superoscillations in the transverse plane of an optical field are relevant in focusing and superresolution. By the close-packing of zeros, one can create localized spots of arbitrarily small width. However, the intensity of such spots will decrease along with the width, and be accompanied by the bright sidelobes characteristic of superoscillations.

3 Subwavelength focusing

In the application of superoscillations to imaging, we note that there are two distinct but related problems to consider. The first of these is the creation of a subwavelength bright spot, or focus, in a non-evanescent regime through the use of superoscillations; the second is the development of techniques to use such spots to extract information about an image with resolution exceeding the classical diffraction limit. Here, we consider the first problem and defer the more subtle question of superresolution to the next section.

Through an analogy between quantum revivals and the Talbot effect, in 2006, Berry and Popescu [15] noted that superoscillations can appear when the light is diffracted by a periodic grating, under appropriate conditions. In short, a grating can have its structure and period \( d \) chosen to mimic the function \( g(x) \) given by Eq. (13), but without any evanescent waves contributing to the pattern; the field is, therefore, superoscillatory immediately after passing through the grating. However, an infinite periodic grating produces perfect images of itself at multiples of the Talbot distance \( z_T = 2d^2/\lambda \), which means that the superoscillatory pattern at the grating can be reproduced at propagation distances that are multiple wavelengths away. This perfect imaging breaks down outside of the paraxial approximation, but Berry and Popescu showed numerically that superoscillations, nevertheless, persist over long propagation distances.

In order to produce a superoscillatory spot, which is localized in a two-dimensional (2D) transverse plane, one would analogously expect that one needs to choose a grating for which Talbot-like effects are strong. In 2007, Huang et al. [28] considered the diffraction of light by a quasicrystalline array of nanoholes in an aluminum film. The array was designed to have an approximate tenfold symmetry, and consisted of 14,000 holes in total.

The quasicrystalline array is used because it has long been known [29] that Talbot imaging is optimized when all spatial frequencies of the source field lie upon a circle in \( k \)-space centered on the origin. The reciprocal lattices of quasicrystals do not perfectly satisfy this condition, but have \( N \)-fold rotational symmetry about the origin, with sets of bright spots located on circles of varying radii, making them better candidates than rectangular arrays.

Using light with wavelength \( \lambda = 660 \) nm, the diffraction pattern was measured approximately 5 \( \mu \)m from the array, well outside of the regime where evanescent waves would play any role. A pattern of isolated, extremely narrow bright spots was observed. A scanning near-field optical microscope (SNOM) was used to map the field intensity, and elliptical spots of a size roughly \( 0.54 \lambda \times 0.30 \lambda \) were observed.

The use of an SNOM tip limits the usefulness of this first approach, but soon after, Huang et al. demonstrated [30] that the superoscillatory can be directly imaged by a microscope and mapped into the far-field. Several patterns are shown in Figure 7. At certain distances, extremely

---

Figure 7: The photonic carpets at wavelength \( \lambda = 500 \) nm seen with a conventional microscope, over an area \( 20 \times 20 \) \( \mu \)m\(^2\). The distances are (A) \( z = 7 \) \( \mu \)m and (B) \( z = 8.5 \) \( \mu \)m. Insets show the superoscillatory hotspots. After [30].
narrow low intensity spots surrounded by a high-intensity halo, characteristic of superoscillations, can be seen.

A quite different strategy for creating a subwavelength focal spot was introduced by Wong and Eleftheriades in 2010 [31], using Schelkunoff’s superdirective antenna design as a basis. Instead of starting with a discrete set of sources in the spatial domain, they instead assumed a discrete set in the spatial frequency domain, of the form

$$
\hat{U}_{\text{src}}(k_x) = \sum_{n=0}^{N-1} b_n \delta(k_x - n\Delta k + k_0),
$$

where $k_x$ is the transverse wavenumber, $\Delta k = 2k_0/N$ and $k_0 = k_0(N-1)/N$. This results in $N$ equally-spaced spectral lines in the range of propagating plane waves, with $-k_0 \leq k_x \leq k_0$.

In the image plane, the field is taken to be of the form,

$$
U_{\text{img}}(x) = e^{i2\pi k_0 x} \sum_{n=0}^{N-1} a_n w^n,
$$

with $w = \exp[i\pi \Delta k]$. As in the case of superdirective antennas, the coefficients of $a_n$ may be placed in such a way as to create a subwavelength-size focal spot.

Next, we apply a backpropagation algorithm. First, we determine the image field in the spatial frequency domain,

$$
\hat{U}_{\text{img}}(k_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} U_{\text{img}}(x) e^{-ik_x x} dx = \sum_{n=0}^{N-1} a_n \delta(k_x - n\Delta k + k_0),
$$

already analogous to Eq. (31). As $\hat{U}_{\text{img}}(k_x)$ represents the spectrum of plane waves with transverse wavenumber $k_x$, we may backpropagate it through multiplication by an inverse propagator, so that

$$
\hat{U}_{\text{src}}(k_x) = \sum_{n=0}^{N-1} a_n e^{-ik_x x} \delta(k_x - n\Delta k + k_0),
$$

where $k_z = \sqrt{k_0^2 - k_x^2}$. Therefore, a choice of source coefficients $b_n = a_n \exp[-ik_z x]$ will produce the desired superoscillatory focal spot at distance $z$. The spatial distribution of the source should then be

$$
U_{\text{src}}(x) = \sum_{n=0}^{N-1} a_n e^{i(n\Delta k - k_0) x} e^{-ik_z x},
$$

where $k_z = \sqrt{k_0^2 - (n\Delta k - k_0)^2}$.

The disadvantage of this simple approach is that the field in the source domain is of infinite extent, due to the localization of the spatial frequencies. To demonstrate the effectiveness of their technique, Wong and Eleftheriades instead considered such an approach for modes in a 2D metal waveguide. If we take the width of the waveguide to be $d$, and treat the waveguide as having infinite conductivity, then we arrive at the following boundary condition for symmetric waveguide modes:

$$
\cos\left(\frac{k_d x}{2}\right) = 0,
$$

or a discrete spectrum of modes with

$$
k_x = \frac{(2n+1)\pi}{d}.
$$

Let us consider the case where $d = 2\lambda$; then only the $n=0$, $n=1$ symmetric modes appear. The $n=0$ mode only has zeros on the boundary of the waveguide; the $n=1$ mode has an additional set of zeros. By choosing the weights of these two modes appropriately, we can place these zeros as close to the center axis of the beam as possible.

As a simple example, we take the modes to be simple cosine functions, and choose weights $a_0 = -0.85$, $a_1 = 1$. Figure 8 shows the intensity of the total field in the image plane and the source plane. The intensity of the field

![Figure 8](image_url)

**Figure 8:** The superoscillatory function designed by the Wong and Eleftheriades method, showing both the intensity in the image plane and the source plane.

The waveguide is taken to have $d = 2\lambda$, and $z = 5\lambda$. The dashed line indicates the oscillation of the highest propagating mode of the waveguide.
in the image plane, compared with the intensity of the \( n = 1 \) mode, shows that we do, in fact, have a superoscillation in the center of the pattern.

With only a small number of propagating modes excited in the waveguide, it is a relatively straightforward exercise to design a source that provides the proper weights. In both simulations and later experiments [32], Wong and Eleftheriades used a waveguide with \( d = 3\lambda \), and used 5 line sources to generate the appropriate excitation; excellent agreement between theory and experiment was found.

As the Wong-Eleftheriades method shows, it is possible to design source fields that produce superoscillations at multi-wavelength distances from the source plane. The most natural, and now common, method to produce superoscillatory spots follows from this observation, and is the construction of superoscillatory lenses (SOLs), which modulate the amplitude and phase of an incident field to produce a subwavelength focal spot.

The first design of a superoscillatory lens was a one-dimensional lens introduced by Huang and Zheludev in 2007 [33], using the prolate spheroidal wavefunctions as a basis; the design principle may be summarized as follows. Given that superoscillations are known to have strong sidebands, one can create a hotspot of prescribed shape in the regime \(-\Delta x \leq x \leq \Delta x\). The bandwidth of the field is naturally limited to only propagating fields, \(-\Delta k \leq k \leq \Delta k\). We may choose to expand the target wavefield using PSWFs with those constraints, with \( c = \Delta x \Delta k\); the PSWFs are naturally orthogonal over the finite \( x \) domain. Therefore, we write

\[
U_{\text{src}}(x) = \sum_{n=0}^{N} a_{n} \psi_{n}(c, x),
\]

where the coefficients \( a_{n} \) can be found from the orthogonality relations between the PSWFs \( \psi_{n} \), \( c, x \). Then, as in the aforementioned Wong-Eleftheriades method, we convert this to a spatial frequency spectrum via Fourier transform

\[
\tilde{\psi}(c, k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(c, x)e^{-ikx}dx = (-i)^{n} \sqrt{\frac{\Delta x}{2\pi \Delta k \lambda_{n}}} \psi_{n}(c, \Delta xk_{n} / \Delta k).
\]

Therefore, we have

\[
\tilde{U}_{\text{src}}(k) = \sum_{n=0}^{N} a_{n} (-i)^{n} \sqrt{\frac{\Delta x}{2\pi \Delta k \lambda_{n}}} \psi_{n}(c, \Delta xk_{n} / \Delta k).
\]

Finally, the source field \( U_{\text{src}}(x) \) can be found by taking the inverse Fourier transform of the above. If the illuminating field is a normally incident plane wave, then the transmission function \( t(x) \) should be chosen as \( t(x) = U_{\text{src}}(x) \).

The design of SOLs that operate in a 3D space is much more involved. In 2012, Rogers et al. introduced a radial binary mask that can create a superoscillatory hotspot surrounded by a high-intensity ring [34]; the details of its use in imaging are described in the next section. The lens was designed using an evolutionary algorithm known as the binary particle swarm optimization [35], which resulted in a lens of 25 transparent regions. Then, under illumination by a laser with \( \lambda = 640 \) nm, the mask generated a focal hotspot with a diameter of 185 nm at a distance of 10.3 \( \mu \)m.

An SOL designed by the previous strategy will produce a hotspot at a very specific propagation distance, and it will be surrounded closely by intense sidebands that limit its usefulness for applications. Drawing inspiration from the Arago spot formed behind a uniformly-illuminated opaque disk, Rogers et al. designed [36] an SOL with an obstructed center; the result is a “super-oscillatory optical needle,” which has an extended superoscillatory brightspot that extends over some 5 \( \mu \)m along the propagation distance. Furthermore, the sidebands of this image are pushed much further away from the central peak; the tradeoff is that the peak width increased to 0.42 \( \lambda \) from 0.35 \( \lambda \) for the traditional SOL. The point spread function of this optical needle lens was analyzed in detail by Roy et al. [37]. Simulations of an optical needle for various blocking regions are shown in Figure 9, which indicates that, as the size of the blocking region increases, the length of the needle increases but the intensity decreases.

Another method for developing subwavelength spots via an optimization process was introduced in 2011 by Mazilu et al. [38], called the optical eigenmode method, and the narrow peaks found can be seen to be the result of superoscillation. We consider the superposition of \( N \) modes \( U_{n}(r) \) taken from some preselected set of beams, such as Laguerre-Gauss beams or Bessel beams, of the form

\[
U_{\text{tot}}(r) = \sum_{n=0}^{N} U_{n}(r).
\]

We are interested in finding the narrowest peak within a windowed region of interest \( S \) near the origin, and we use the variance in this region as the measure, defined as
G. Gbur: Superoscillations for imaging

\[ \Delta^2 = \frac{1}{2} \int_0^\lambda r^2 U_n^* (r) U_{n1} (r) dr \]

On substituting the expression for the total field, we have

\[ \Delta^2 = 2\pi \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} a_n a_m \int_0^\lambda r U_n^* (r) U_m (r) r dr \]

where, for simplicity, we assume we are using radially symmetric modes, have integrated out the azimuthal dependence, and will choose as our region of interest a region of radius \( \lambda \). The expression for \( \Delta^2 \) may be written in a vector-matrix form as

\[ \Delta^2 = \frac{a^H M^{(2)} a}{a^H M^{(0)} a} \]

where \( |a\rangle \) is a vector of the elements \( a_n \) and \( M^{(2)} \) is a Hermitian matrix with elements,

\[ M^{(2)}_{nm} = \int_0^\lambda r U_n^* (r) U_m (r) r dr \]

and

\[ M^{(0)}_{nm} = \int_0^\lambda U_n^* (r) U_m (r) rdr \]

Ideally, we would like to reduce the minimization problem to finding the smallest eigenvalue of the matrix \( M^{(0)} \); however, because our modes are not orthogonal in the region of interest, as indicated by the matrix \( M^{(0)} \) being non-diagonal, the denominator will significantly effect the value. Therefore, we introduce a pair of matrices into our expression: the first matrix \( B \) will diagonalize \( M^{(0)} \) via a similarity transformation, while the second matrix \( D \) is a diagonal matrix with elements equal to the square roots of the inverses of the eigenvalues of \( M^{(0)} \). The matrix \( D^{-1} \) is a diagonal matrix with elements equal to the square roots of the eigenvalues of \( M^{(0)} \). Then, we have

\[ \Delta^2 = \frac{a^H B^H D^H M^{(2)} D B^H a}{a^H B^H D^H M^{(0)} D B^H a} \]

We now let \( |b\rangle = D^{-1} B |a\rangle \), which simplifies the previous expression to

\[ \Delta^2 = \frac{b^H M^{(2)} D^H b}{b^H D^H M^{(0)} D b} \]

In this form, the eigenvector \( |b\rangle \) of the combined matrix with the smallest eigenvalue will also produce the smallest spot in the region of interest. To perform this minimization procedure, then, we take the following steps: 1. generate \( M^{(0)} \) and \( M^{(2)} \); 2. use \( M^{(0)} \) to determine the matrices \( B \) and \( D \); 3. generate the matrix \( N = DBM^{(2)} D^H B^H \); 4. find the smallest eigenvalue of \( N \); 5. find the corresponding vector \( |b_{\text{min}}\rangle \) and 6. find \( |a_{\text{min}}\rangle = B^H |b_{\text{min}}\rangle \).

An illustration of this process is shown for Laguerre-Gauss beams \( LG_{n0} \) in the waist plane of width \( w_0 = 2\lambda \) in Figure 10; the region of interest is taken to be a circle of radius \( \lambda \).
radius $\lambda$. One can clearly see the reduction in spot size as the number of Laguerre-Gauss modes is increased; however, this reduction comes with a huge increase in the intensity of a halo ring outside of the region of interest, which is characteristic of superoscillations.

This technique was demonstrated experimentally by Baumgartl et al. [39] in 2011, using a set of 15 Bessel beams as the basis set. An SLM was used to generate the desired annular spatial frequencies of the Bessel beams, which was then imaged into the focal plane of a microscope objective. The subwavelength patterns were imaged using an SNOM tip. Excellent agreement with the theoretical model was found.

The superoscillatory focusing described so far was done in the transverse plane. It is also possible to take an appropriate superposition of fields with different wavenumbers in the $z$-direction to produce a subwavelength superoscillatory spot along the axial direction. This was demonstrated by Zacharias et al. [40] in 2017, using a set of Bessel beams of identical order $m = 0$ but with different axial wavenumbers $\beta$, as defined later in Eq. (58) in Section 5.

Given that it is possible to reduce the spot size in the transverse plane or in the axial direction, it also possible to do both simultaneously and make, for example, a spherical subwavelength spot. To do so, one would need to superimpose multiple radially symmetric waves, in a combination of the form

$$U(r) = \sum_{n=1}^{N} f_n(r).$$

However, for monochromatic waves, this is not possible: all propagating waves have their wavevectors constrained to a spherical surface, such that $k_x^2 + k_y^2 = k_0^2$, and the only radially symmetric superposition of these modes has the form $U(r) = \sin(k_0 r) / k_0 r$. However, if multiple frequencies are used, i.e. a superoscillatory field is constructed in the time domain, then a superoscillatory field can be constructed of the form

$$u(r, t) = \sum_{n=1}^{N} A_n e^{i \omega_n t} \frac{\sin(k_n r)}{k_n r},$$

where $k_n = \omega_n / c$. For a given time $t = t_0$, one can design the overall field pattern to be superoscillatory.

This approach was theoretically investigated by Wong and Eleftheriades [41] in 2017, in which they demonstrated a 3D spot size roughly 66% of the diffraction limit. However, this spot remains subwavelength for only a fraction of the overall waveform period, indicating that more sophisticated techniques are required for image reconstruction. One possibility is to measure both the amplitude and phase of the light scattered in the imaging process, from which the scattered field can be in principle reconstructed during the superoscillation duration.

Superoscillatory lenses, like diffractive lenses in general, tend to be strongly wavelength-dependent and, therefore, suffer from chromatic aberrations. Recently, Yuan, Rogers and Zheludev [42] introduced two algorithmic strategies for making achromatic superoscillatory lenses. As the focus of a SOL lens can be stretched in the longitudinal direction, one possibility is to design the lens so that the extended foci at multiple wavelengths overlap, thus providing a smaller region where all wavelengths are focused simultaneously. Given that SOL lenses also tend to produce multiple hotspots along the longitudinal direction, another possibility is to make hotspots of different orders overlap for different wavelengths at a single focus.

Several other demonstrations of superoscillatory focusing and their applications are worth noting. In 2015, David et al. [43] introduced planar waveguide platforms that can be used to focus visible light on the nanoscale; they designed one to produce a superoscillatory focal spot as an example of the technique’s flexibility. Superoscillatory focusing has also been used in optical trapping and manipulation. For example, Singh et al. [44] used a superoscillatory beam to trap and manipulate polystyrene particles.

### 4 Superresolution imaging

The remarkable ability to generate subwavelength spots in a bandlimited signal comes with significant challenges, especially in the context of imaging. A superoscillatory spot is necessarily surrounded by a ring-shaped high intensity halo, which can overwhelm the subwavelength signal and restricts the usable imaging area, as we will see. An additional discussion of the limitations of superoscillation in microscopy can be found in Hyvärinen et al. [45]. Furthermore, because superoscillation is the result of a delicate interference effect between waves, this sideband cannot be windowed away, as the superoscillations will be destroyed on further propagation. From a practical perspective, the fabrication of an SOL requires very precise manufacturing on the nanometer scale in order to create the focal spot.

Despite these obstacles, much progress has been made, both theoretically and experimentally, in the use of superoscillations to break the conventional diffraction limit in imaging. In the simplest realization, as described...
theoretically by Huang and Zheludev [33], one can image closely-spaced incoherent objects, provided they lie within each other’s halos.

To illustrate this, we evaluate a simple simulation of an SOL in 2D space, and compare its performance to an ordinary lens under the same circumstances. We consider a lens that lies within an aperture of total width \( a \), which will take an object at a rear plane at distance \( z = -d_o \) and image it at a front plane at \( z = +d_i \). The Fourier theory indicates that the image produced is given by the Fourier transform of the transmission function of the lens; for an ordinary lens, this transmission function is constant, and the image is given by

\[
\Delta \mathbf{U}_x = \frac{\lambda}{a} \text{sinc} \left( \frac{k \Delta}{2} \right),
\]

where \( \Delta = \frac{x}{d_o} + \frac{x}{d_i} \) with \( x \) as the position of the point object. The minimum separation resolvable between two points is given by the distance to the first zero of the sinc function, or \( k \Delta_{\text{min}} = 2\pi/a \).

We assume that the SOL has the same lensing effect as the ordinary lens, but also has a superoscillatory transmission function in the aperture. We take the field in the image plane as

\[
\mathbf{U}_x = \left[ (k \Delta)^2 - b^2 \right] \left[ (k \Delta)^2 - c^2 \right] \hat{f}(k \Delta),
\]

where \( f(x) \) is a function that is bandlimited in space to \( |x| \leq a/2 \). Following the Chremmos and Fikioris method, we use

\[
f(x) = \cos^8(\pi x/a),
\]

and we take the positions of the zeros to be \( b = 0.5(2\pi/a) \), \( c = 0.9(2\pi/a) \), which places one pair of zeros at half the lens resolution distance; the other pair of zeros is taken to push the halo further from the imaging region. The Fourier transform of Eq. (54) is readily determined, from which one can obtain the superoscillatory field pattern from Eq. (53).

Next, we compare the incoherent imaging of two point sources via a traditional lens and our SOL in Figure 11. The two point sources are clearly distinguishable for the SOL, but not for the traditional lens. The high-intensity halo around the image is clearly seen as well.

To image larger areas, a confocal scanning geometry must typically be implemented. This was first done by Kosmeier et al. [46], using their optical eigenmode method [38] to generate the superoscillatory beams. In their experiment, the superoscillatory spot was scanned transversely across a sample consisting of two pinholes in an opaque screen. At the detector, a small \( 3 \times 3 \) pixel spot moved with the beam to act as a virtual pinhole to filter out noise from the spot halo. Working with a He-Ne laser, the researchers were able to increase the resolution by a factor of 1.36 relative to the diffraction-limited case.

A scanning geometry was also implemented by Rogers et al. to achieve superresolution [34], with an SOL providing the subwavelength focal spot. Single and double slits were imaged as well as closely-spaced nanoholes in a metal film; using a diode laser with \( \lambda = 640 \) nm, they were able to reliably resolve holes separated by \( \lambda/6 \). A comparison of the results with an ordinary lens and the SOL is shown in Figure 12.

It should be noted that a superoscillatory spot is typically elongated along the direction of propagation; as noted by Wong and Eleftheriades [41], it is not generally possible to make a 3D superoscillatory spot with monochromatic light. This elongation means that such superoscillatory scans are expected to be effective only for thin samples.

When coherent imaging data are available, it is possible to use superoscillatory post-processing techniques to construct a superresolved image of the object, as Amineh

![Figure 11](https://example.com/figure11.png)

**Figure 11**: The comparison of the imaging of two subwavelength-separated sources via (A) an SOL lens and (B) a traditional lens. Here \( d_o = d_i = 30 \) mm, \( a = 10 \) mm, \( \lambda = 500 \) nm, and the sources were separated by \( 2/3 \)rds the resolution limit. The dashed line indicates the total image.
and Eleftheriades demonstrated [47]. To understand this, let us suppose that the object field in plane \( z = 0 \) is \( U_o(x) \). This function has a Fourier transform \( \tilde{U}_o(k) \). The propagation of the field to the image plane at \( z \) can be done using the expression

\[
\tilde{U}_i(k, z) = \tilde{U}_o(k) e^{ikz},
\]

(55)

where \( k = \sqrt{k_0^2 - k_s^2} \).

If the image plane is more than a few wavelengths away from the object plane, then only the propagating waves will survive, with \( |k_s| \leq k_0 \). If the image field is measured via holography, then, in principle one can determine the function \( \tilde{U}_i(k) \) by using the inverse propagator \( \exp[-iknz] \), and then take an inverse Fourier transform to get a bandlimited image of the object,

\[
U_o(x) = \int_{-k_0}^{k_0} \tilde{U}_i(k) e^{-ikx} dk_s,
\]

(56)

This image will be resolution-limited due to the absence of high spatial frequencies. However, we can improve the resolution without changing the bandlimit by multiplying the backpropagated Fourier spectrum by a superoscillatory spectrum \( \tilde{f}(k_s) \), i.e.

\[
\tilde{U}_b(k) = \tilde{U}_i(k) e^{-ik\phi z} \tilde{f}(k_s).
\]

(57)

As the Fourier spectrum of a point source at \( x_0 \) is simply proportional to a complex exponential, \( \exp[ik \cdot x] \), the inverse transform of the modified spectrum will be a superoscillatory peak at the point \( x_0 \). Two closely-spaced points will have a reconstructed image of two closely-spaced superoscillatory peaks.

Due to the halo effects, the technique of Amin et al. and Eleftheriades is still limited to resolving very closely-spaced objects in a limited area. It was extended by Li, Li and Cui [48] with additional computational techniques in order to allow it to be used for larger and more general objects.

As has been emphasized repeatedly, a fundamental limitation of superoscillation in imaging is the presence of the bright halo, which increases in intensity relative to the center peak as the width of the center peak becomes more subwavelength. An alternative philosophy for superresolution through superoscillation was recently introduced by Dong et al. [49], focusing on a reduction of the ordinary sidelobes of a point spread function. This modified scheme, the zeros are densely packed to suppress the first sidelobe of the PSF, as this sidelobe also degrades the resolution. As these zeros are not being used to narrow the central peak, the halo surrounding it is of much lower relative intensity than in conventional superoscillation schemes.

Most of the work on superoscillation to this point has focused on scalar wavefields. Very recently, Kozawa, Matsunaga and Sato [50] investigated the structure of radially polarized Laguerre-Gauss modes and demonstrated that such modes tend to naturally form superoscillatory spots as the size of the incident beam is varied. They were able to see spots on the order of 100 nm in size using incident light of 532 nm, and performed improved confocal imaging of HeLa cells with such beams.

Research on superoscillation imaging has now passed the “proof-of-concept” stage and is being optimized with an eye towards more practical implementation. Quite recently, Rogers et al. [51] introduced an algorithm to simultaneously optimize the spot size and intensity relative to the size of the halo for different desired fields of view. Such a tool will allow researchers to readily tailor their superoscillations to their particular application.

5 Structured beam superoscillations

The examples we have seen so far demonstrate that superoscillations can persist or revive over multi-wavelength distances, but it has also been shown that they can be carried over arbitrarily long distances, in principle, using nondiffracting waves.

It has been long known that one can construct diffraction-free beamlike solutions to the scalar wave equation [52, 53], and that such beams have become an important tool in a number of optical applications. The most commonly considered class are Bessel beams, of the form

\[
U_m(r, \phi, z) = J_m(\alpha r)e^{im\phi}e^{ikz},
\]

(58)

where \( r^2 + \beta^2 = k_0^2 \), in which \( \beta \) represents the propagation constant of the beam. From the integral definition of the
Bessel functions, we can deduce that a Bessel beam may be viewed as a coherent superposition of plane waves with their wavevectors aligned on a cone, with transverse wave-number \( c \). Provided \( c^2 < k_j^2 \), these beams are propagating and spatially bandlimited in the transverse plane. As the Bessel function asymptotically decays as \( 1/\sqrt{r} \) for large \( r \), the total energy in an ideal Bessel beam is infinite, making them not realizable in practice. However, the approximate forms of such beams with Gaussian envelopes will have finite energy and can remain approximately diffraction-free over long propagation distances [54].

Superoscillations can arise from an appropriate weighted sum of bandlimited waveforms, and this observation applies to both nondiffracting beams as well as plane waves. As there are an infinite set of Bessel beams with the same propagation constant \( \beta \), it is possible to construct a tailored superposition of such beams that possesses superoscillations, and those superoscillations will propagate without change over a long distance.

This approach was introduced and tested by Makris and Psaltis [55] in 2011. To design their superoscillations, they required that a superposition of \( N \) Bessel beams pass through \( N \) predetermined values at points \( r_m \), with \( m = 0, 1, 2, \ldots, N-1 \). Their field \( g(x, y, z) \), therefore, was of the form,

\[
g(x, y, z) = \sum_{m=0}^{N-1} c_m U_m(r, \phi, z). \tag{59}
\]

A simple example is one in which three points are taken on the \( x \)-axis at \( x_0 = -\delta, x_1 = 0 \) and \( x_2 = +\delta \) as well as \( g(x_0) = 0, g(x_1) = 0, \) and \( g(x_2) = 1 \). Provided \( \delta \) is sufficiently small, the Bessel functions can be approximated by the first term of their Taylor series, \( J_m(x) = x^m/(2^m m!) \), making the solution for the coefficients easier. The result is \( c_0 = 1, c_1 = 0 \) and \( c_2 = -8/(a\delta)^2 \).

The resultant intensity is shown in Figure 13. Though the maximum transverse spatial frequency in the beam is \( a = 2 \), with a corresponding wavelength of \( 2\pi/a = \pi \), the zeros are spaced by approximately 0.5. It can be seen from Figure 13B that the zeros are optical vortices, with a characteristic circulation of the phase around their cores.

A modified version of this scheme was introduced and tested experimentally by Greenfield et al. in 2013 [56]. They generalized the superposition to include not only all orders of Bessel beams, but lateral shifts of each beam, in the form

\[
g(x, y, z) = \sum_{m=0}^{N} \sum_{l=0}^{l_m} a_{lm} U_m(r - r_m), \tag{60}
\]

where there are \( l_m \) beams superimposed for every order \( m \) included in the superposition, and \( r_m \) is the shifted origin of the beam with indices \( l \) and \( m \).

The use of shifted beams allows the superoscillations to be generated experimentally using shifted copies of a single order of Bessel beam. In their experiments, Greenfield et al. used essentially a Mach-Zehnder interferometer to produce two copies of a second-order Bessel beam, with an attenuator in one arm of the interferometer producing the proper amplitude ratio and a 3D stage in the other arm producing the lateral shift. Each arm contained its own spiral phase plate to produce a second-order vortex; after being recombined, the two beams were passed together through an axicron to produce the Bessel profile. Their experimental results are the first demonstration of superoscillations in a nondiffracting beam.

The results for nondiffracting beams discussed so far are in the scalar approximation. In 2016, however, Makris et al. [57] theoretically constructed superoscillations from superpositions of vectorial electromagnetic Bessel beams.

Another class of nondiffracting beams are the Airy beams, which were first introduced in the context of quantum mechanics in 1979 [58] and were much more recently generated optically [59]. Such beams have an
amplitude proportional to an Airy function along one axis in a source plane, i.e.

\[ U(x, 0) = \text{Ai}(x / x_0) \],

(61)

with \( x_0 \) as a length parameter. As illustrated in Figure 14, not only are such Airy beams nondiffracting but they also demonstrate self-acceleration on propagation. Like their Bessel counterparts, the idealized Airy beams have infinite energy, but they can be realized by adding an exponential envelope to their slowly-decaying section.

It is also possible, through a very clever construction introduced by Eliezer and Bahabad [60], to introduce superoscillations into Airy beams by an appropriate superposition. Looking at the left side of the Airy beam in Eq. (14), it is to be noted that the oscillations are essentially sinusoidal. By comparing a superposition of Airy patterns with a superposition of sinusoids, one can match the Airy superposition to a known superoscillatory pattern. In particular, we return to the Berry Popescu model of superoscillations, Eq. (13), and note that it can be expanded as a sum of complex exponentials as

\[
g(x) = \left( \frac{a + 1}{2} \right)^N \sum_{m=0}^{N} \binom{N}{m} (-1)^m \left( \frac{a + 1}{a - 1} \right)^m e^{i(\pi/4)(N-2m)}. \]

(62)

This can be proven by taking the original form of \( g(x) \), writing the sine and cosine in terms of complex exponentials and doing a binomial expansion with respect to the terms \( \exp[ix] \) and \( \exp[-ix] \). The imaginary part of this function will be a simple sum of sines. In order to construct a superoscillating Airy function, one takes a superposition of Airy functions with different widths

\[ f(x) = \sum_{i=0}^{2N} A_i \text{Ai}(x / \alpha_i) \exp(yx), \]

(63)

where \( y \) is a common exponential decay factor. Noting that, asymptotically, Airy functions may be written in the form

\[
\text{Ai}(x) = \frac{x^{-1/4}}{\sqrt{\pi}} \sin \left[ \frac{2}{3} x^{3/2} + \pi/4 \right],
\]

(64)

one can substitute from this expression into Eq. (63). Then, by choosing the weights of Eq. (63) to make it match the imaginary part of Eq. (62), one can construct an Airy function that exhibits superoscillatory behavior.

An example of such oscillations is shown in Figure 15. For this particular case, it can be seen that there is modest enhancement of the local frequency in the area where the intensity is low, as one would expect from superoscillations. The superoscillations derived from this method are limited to modest gains, though. In order for Eq. (63) to match Eq. (62), the common \( \pi/4 \) phase shift of the asymptotic Airy functions must be small compared to the period of the local frequency of superoscillation, or \( \pi/4 < 2\pi/aN \). Nevertheless, the model was confirmed by experiments, using a spatial light modulator to produce the Airy pattern.

6 Other advances in superoscillations

In this section, we consider a number of recent advances in the study of superoscillations that have much practical potential. Though not directly related to imaging,
they illustrate the possibilities inherent in the study of superoscillations.

In 2017, Eliezer and Bahabad demonstrated [61] that it is possible to create sub-oscillations, an effect complimentary to that of superoscillations. A sub-oscillation may be defined as an oscillation that appears in a function, which has a minimum frequency in its bandwidth, i.e. if the function is bandlimited to $\Delta \omega \leq |\omega|$, then a sub-oscillation is an oscillation that occurs with a frequency $|\omega| < \Delta \omega$. The authors first demonstrated this effect by considering the superoscillation model of Eq. (7), namely,

$$g(x) = \frac{1}{\delta N 2\pi} \int_{-\infty}^{\infty} e^{i(x+\delta u)} \exp \left[ -\frac{1}{2\delta^2}(u-\mu)^2 \right] du,$$

but chose $k(u) = \cosh(u)$, which has a lower frequency bound of $u=1$. Again by an asymptotic calculation, one can demonstrate that the local oscillations for small $\delta$ will have the approximate wavenumber

$$k(\mu) = \cosh(\mu) = \cos(\mu),$$

which satisfies $|k(\mu)| < 1$; the function is sub-oscillatory. The familiar Berry Popescu model of Eq. (13) can also be modified to demonstrate sub-oscillations by taking its inverse, i.e.

$$h(x) = \frac{1}{\cosh(x)+ias\sin(x)}.$$

The function $h(x)$ can be shown to have sub-oscillations in its local frequency if $0 < a < 1$.

The existence of sub-oscillations can be readily deduced from the aforementioned method of Chremmos and Fikioris [21] as well. Just as the zeros of a function can be placed closer together without changing the bandwidth, producing superoscillations, the zeros can be spaced further apart to produce the opposite effect. Just as in the superoscillation case, the amplitude of the field will be dramatically lower in the regime of sub-oscillations.

Eliezer and Bahabad used the existence of suboscillations to experimentally demonstrate super defocusing. The scenario they envision is a desire to have a beam that spreads as rapidly as possible on propagation; in the simplest case, this can be achieved by first tightly focusing the beam. The angular spreading of the beam and the distribution in the focal plane are related by a Fourier transform, which means that, typically, a beam with a broad spreading will have a small focal spot. However, if a small obstruction is placed just before the focus, it puts an upper limit on the spot size at focus; if it is too small all the energy will be blocked. Through the use of sub-oscillations, however, it is possible to “stretch” out the focal spot transversely, allowing energy to pass around the obstruction, without reducing the spread in the far-field.

Suboscillations, at first glance, may seem less spectacular than their superoscillating cousins, but they can also produce seemingly paradoxical effects. Chremmos, Chen and Fikioris [62] introduced their own method of designing suboscillations, and were able to demonstrate the possibility of using suboscillations to approximate arbitrarily well any function within a finite interval using them, including a constant function.

Returning to the topic of superoscillations, it is to be noted that most of the investigations in optics have considered the effects of the space/spatial frequency domain, not on the time/frequency domain. Of course, if it is possible to improve the resolution in the spatial domain through the use of superoscillations, it should be similarly possible to improve temporal resolution. This problem was studied in the context of pulse compression by Wong and Eleftheriades [63] in 2011, who once again used the Schelkunoff approach to design their waveforms. In this case, they considered pulses formed from a finite number $N$ of equally spaced spectral lines; the total field can then be written as a polynomial in the form

$$g(t) = e^{i\omega_0 t} \sum_{n=0}^{N-1} a_n z^n,$$

where $z = \exp[i\Delta \omega t]$. This expression is then analogous to Schelkunoff’s expression for antenna arrays, and the same techniques can be applied to derive superoscillatory pulses. The results were confirmed by experiment.

Another approach to superoscillating pulses was taken some years later by Eliezer et al. in 2017 [64]. They return again to the function of Berry and Popescu in the form,

$$g(t) = \left[ \cos(\Omega_0 t) + i\sin(\Omega_0 t) \right]^N,$$

which has a fundamental frequency $\Omega_0$, and note that the real part may be written as

$$\text{Re}\{g(t)\} = \sum_{n=0}^{[N/2]} A_n \cos(q_n \Omega_0 t),$$

where $[N/2]$ represents the floor of $N/2$ and $q_n = 2\pi n + \mu_n$, with $\mu_n = \text{mod}(N, 2)$ representing the modulus of $N/2$. This representation of $\text{Re}\{g(t)\}$ can be derived from $g(x)$ in Eq. (62), though it takes some effort; here we simply note that they are clearly structurally similar.
The limitation of Eq. (70) is that it does not look like the spectrum of a pulse, which typically has a finite bandwidth \( \Delta \omega \) and a central frequency \( \omega_c \); Eq. (70), in contrast, even contains the frequency \( \omega = 0 \) if \( N \) is even. The solution is to derive the superoscillations as beats superimposed on a carrier signal with central frequency \( \omega_c \).

As an example, Eliezer et al. constructed a superoscillating signal with \( N = 3 \) and \( a = 2 \), producing a \( g(t) \) of the form

\[
g(t) = -\frac{9}{4} \cos(\Omega_o t) + \frac{13}{4} \cos(3\Omega_o t). \tag{71}
\]

They combined four beat signals with frequencies \( \nu_n = \{370.9, 373.2, 375.6, 377.9\} \) THz, and appropriate amplitude and phase relations taken to match Eq. (70). The signal is modulated with a Gaussian envelope of the form \( \exp[-t^2/2\sigma^2] \), with \( \sigma = 280 \) fs. This envelope broadens the overall bandwidth of the signal but does not significantly change the nature of the superoscillations. An illustration of the superoscillation intensity and a Gaussian of the same width are shown in Figure 16; as can be seen, there is a small superoscillatory spot in the center of the pulse.

This model of a superoscillatory pulse was confirmed experimentally using a femtosecond laser source with a 4\( f \) pulse shaper to produce the pulse and a FROG apparatus to measure the resulting signal.

A collaboration from some of the same researchers recently introduced a more elegant method for producing simple superoscillations in a pulse [66], using a technique first applied in the spatial domain [44]. In this method, one shifts the spectral envelope \( \tilde{f}(\omega) \) of a pulse by a phase of \( \pi \) over a finite range of frequencies, that is, introducing the function

\[
\tilde{g}(\omega) = \begin{cases} 
1, & |\omega| < \omega_s, \\
0, & \text{otherwise,}
\end{cases} \tag{72}
\]

we introduce a modified function \( \tilde{g}(\omega) \) as

\[
\tilde{g}(\omega) = \tilde{f}(\omega) \left[ 1 - 2\tilde{s}(\omega) \right]. \tag{73}
\]

The creation of superoscillatory zeros can be understood as arising from the relatively narrow function \( f(t) \) interfering with the broader pulse created by the narrower band function \( -2\tilde{f}(\omega)\tilde{s}(\omega) \). The result is a small superoscillatory spot, bounded by a pair of zeros, in the middle of the pulse. This spot first decreases in size as \( \omega_s \) increases, but eventually disappears as \( \omega_s \) becomes comparable to the original bandwidth of the function \( f(t) \). The theoretical predictions were again confirmed with FROG measurements.

Notably, in the time domain, superoscillations have the potential to overcome other limitations of light propagation. In 2014, Eliezer and Bahabad considered [66] the propagation of superoscillations with an effective frequency \( \omega_s \) through an absorptive medium with a resonance at that frequency. As the actual spectrum of the pulse does not overlap the resonance frequency, in principle, it is possible to propagate a superoscillatory signal at \( \omega_s \) through a medium that would be opaque to normal oscillations. Given that the superoscillations are a delicate interference effect, one natural concern is that ordinary dispersion effects would destroy them after a short propagation distance. However, it was found that the superoscillations underwent semi-periodic revivals in the medium, suggesting that they can be transmitted over significant distances under the right circumstances.

One other remarkable property of superoscillations in the time domain is worth noting here. One of the earliest works addressing the implications of superoscillations was written, but unpublished, by Aharonov, Popescu and Rohrlich in 1990 [67], in which they suggested that a box filled entirely with red light could possibly release gamma radiation when a window was opened. If the window is opened only over a short period of time while superoscillations are incident upon it, those oscillations of arbitrarily high frequency will escape the box, and will evidently maintain their high-energy nature. In 2018, Berry and Fishman [68] performed a detailed theoretical analysis of the problem, and confirmed that this prediction, as paradoxical as it might seem, is correct. The “missing” energy needed to convert red light into gamma radiation comes from the opening and closing of the shutter over a very short timescale. Therefore, like many paradoxes of physics, the strange nature of superoscillations turns out to fit beautifully into our existing knowledge.
7 Concluding remarks

For over a century, optical scientists largely viewed the resolution of imaging systems to be limited to roughly a half-wavelength, based on the criteria of Rayleigh and Abbe. In Abbe’s case, for example [69], the minimum spacing $r_{\text{min}}$ between two imaged points was defined as

$$r_{\text{min}} = 0.6098 \frac{\lambda}{NA},$$

(74)

where $\lambda$ is the wavelength of light and $NA$ is the numerical aperture. In the 1920s, Synge pointed out that measurements in the near-field of an object can allow sub-wavelength resolution [70], but this strategy was largely ignored until the 1980s, when Pohl, Denk and Lanz produced the first optical near-field images [71]. Although this removed the classical resolution limit as a barrier to imaging, new limitations were found, such as the low amount of light collected and the difficulty of interpreting the detected plane wave/evanescent wave signals.

Analogously, though the negative refractive index “superlens” introduced by Pendry in 2000 [72] is theoretically perfect, practical considerations, such as the need to work in the near-field of the lens, the inherent material losses of the lens and the high intensities needed inside the lens [73], limit its effectiveness.

We have already seen in this review that superresolution through superoscillations has its own inherent limitations, such as the instability of the interference effect, the low intensity of the superoscillations and the size of the sidelobes. We do not mention the limitations of near-field-optics and superlensing as a pessimistic view, however, but rather the opposite: both of these superresolved imaging techniques have proven useful and remain vibrant areas of research. Superoscillations offer additional advantages over many existing superresolution techniques, such as the fact that it needs no post-processing of data taken, or prior knowledge of the imaged sample.

Although the investigations of superoscillations in imaging are still quite new and there is much to be done, it is not unreasonable to expect that it will take its place among other superresolution techniques to “bend,” though not completely break, the classical limits on imaging.

Acknowledgement: Greg Gbur is funded by the Air Force Office of Scientific Research under grant FA9550-16-1-0240.

References

[18] Lanz produced the first optical near-field images [71].


