Planar Diffractive Lenses: Fundamentals, Functionalities, and Applications

Kun Huang, Fei Qin, Hong Liu, Huapeng Ye, Cheng-Wei Qiu, Minghui Hong, Boris Luk’yanchuk, and Jinghua Teng*

Improving its image quality and resolution has always been one of the most valuable topics since its invention. Due to the fundamental laws governed by diffraction, the spatial resolution of optical microscopy is restricted to $\approx 200$ nm at visible light, which has been clearly elaborated by Abbe[1] and Rayleigh[2,3] through the classic wave theory of light, or doubly predicted by the uncertainty relationship in quantum mechanics.[4] With the rapid development of nanotechnology, the limited resolution for an optical microscope becomes insufficient in various applications such as biological imaging, materials science, and nanolithography. In order to surpass the diffraction limit and see finer details of objects, tremendous efforts have been put over several decades in developing various superresolution techniques. We assigned these techniques into three main categories: near field, far-field fluorescence-based, and far-field label-free methods.

1. Introduction

Nowadays, optical microscopy is almost ubiquitous in all research fields and industrial sectors for imaging small objects.

1.1. Near-Field Superresolution Microscopy

In Fourier optics, an object can be mathematically taken as a weighted superposition of many spatial frequencies. Upon

Traditional objective lenses in modern microscopy, based on the refraction of light, are restricted by the Rayleigh diffraction limit. The existing methods to overcome this limit can be categorized into near-field (e.g., scanning near-field optical microscopy, superlens, microsphere lens) and far-field (e.g., stimulated emission depletion microscopy, photoactivated localization microscopy, stochastic optical reconstruction microscopy) approaches. However, they either operate in the challenging near-field mode or there is the need to label samples in biology. Recently, through manipulation of the diffraction of light with binary masks or gradient metasurfaces, some miniaturized and planar lenses have been reported with intriguing functionalities such as ultrahigh numerical aperture, large depth of focus, and subdiffraction-limit focusing in far-field, which provides a viable solution for the label-free superresolution imaging. Here, the recent advances in planar diffractive lenses (PDLs) are reviewed from a united theoretical account on diffraction-based focusing optics, and the underlying physics of nanofocusing via constructive or destructive interference is revealed. Various approaches of realizing PDLs are introduced in terms of their unique performances and interpreted by using optical aberration theory. Furthermore, a detailed tutorial about applying these planar lenses in nanoimaging is provided, followed by an outlook regarding future development toward practical applications.
illumination, the scattering light from the object contains propagating waves with low \((< 1/\lambda)\) spatial frequencies and confined evanescent waves with high \((> 1/\lambda)\) spatial frequencies. Conventional optical microscopy can only capture the finite object information carried in propagating waves, leading to a limited resolution due to the absence of high spatial frequencies that only exist within one wavelength from the surface of object. Therefore, the efficient exploitation of evanescent waves is the key to near-field approaches for superresolution imaging.[8]

Near-field scanning optical microscopy (NSOM) is a commonly used technique with the help of an aperture or apertureless nanoprobe that can efficiently approach the surface of an object and then scatter the evanescent waves for subsequent collection by an objective lens. The resolution of NSOM is only determined by the diameter of the probe. NSOM was initially conceptualized by Edward Hutchinson Synge in 1928.[7] Limited by the technical challenges, such as fabricating a nanoscale aperture, achieving a sufficiently intense light source and maintaining the aperture in close proximity to the specimen at the nanometer scale, this visionary proposal had not been realized until 1984, when it was demonstrated by a research group in IBM Zurich laboratory.[8] In 1986, they reported an imaging resolution of below 50 nm.[9–11] The NSOM technique began to evolve as a scientifically useful and important instrument, and its imaging resolution of less than 10 nm can be achieved by aperture-less NSOM.[12]

Using a superlens is another approach to realize superresolution imaging and nanolithography. In 2000, John B. Pendry reported in his epochal paper that a perfect lens enabling the simultaneous collection of both propagating and evanescent waves can be realized by using a material with negative permittivity and permeability.[13] This concept was experimentally demonstrated in 2004 by Luo and Ishihara for realizing optical nanolithography[14,15] and then used to achieve an imaging resolution of 89 nm in terms of a silver superlens in 2005 by Xiang Zhang’s group in Berkeley.[16] Combining with the frequency shift technique, Xiang Zhang’s group reported a far-field superlens that can project a nano-object into a subdiffraction-limit image in 2007.[17,18] The superlens has also found the applications in lithography due to the well-confined evanescent waves.[14,19–22]

In 2011, Wang et al. reported a white-light microsphere-based microscope with an achieved imaging resolution of 50 nm \((\lambda/8)\).[23] Combining with the laser scanning confocal microscope, they demonstrated the enhanced lateral resolution of \(=25\) nm \((\lambda/17)\) under the illumination of 408 nm wavelength.[24] Its working principle is that the dielectric microsphere works as a "photic nanojet"[25] to enhance the near-field illumination as well as convert evanescent waves to magnified propagating waves.[21,24,26,27] The subsequent studies indicated that the imaging performance could be enhanced further by using high-refractive-index microscope, immersing the microsphere by liquid or polymer ambient, vector-beam illumination and decorating the surface of the microscope with engineered structures.[27–31] Recently, a microsphere-based scanning microscopy has been demonstrated with the microsphere attached and fixed onto a cantilever of atomic force microscopy, which is able to realize noninvasive nanoimaging with large angle-of-view at the visible wavelengths.[32]

Due to the near-field excitation of evanescent waves, these approaches face the challenging operation in practice and the potential damage to specimen.

1.2. Far-Field Fluorescence-Based Superresolution Microscopies

The fluorescence technique overcomes the diffraction limit of light by selectively activating and deactivating fluorophores assisted by the optical nonlinear processes. These fluorescence-based microscopes are ideal to study the cell with <200 nm scale in biology, allowing for a deeper understanding about cell structures and tracking activities occurring inside living cells.[33]

Stimulated emission depletion (STED) microscopy was initially proposed by Hell and Wichmann in 1994[34] and experimentally demonstrated later.[35] It works by employing stimulated emission with a doughnut-shape beam to inhibit the fluorescence process at the outmost region of the excited molecules.[36] The doughnut-shape depletion beam and a solid excitation beam are well-aligned in a concentric way. The doughnut-shape beam has longer wavelength than the natural emission photons from the fluorophores. The fluorophores are first excited and then depleted by the doughnut shape beam, leaving an effective fluorescent spot with a subdiffraction-limit size in the nondepleted region. Combining with the confocal
microscopy technique and raster scanning across the specimen, a fluorescence-based superresolution imaging can be achieved. The imaging resolution $\Delta r$ of STED microscopy can be expressed as $\Delta r = \lambda / (2NA \sqrt{1 + I_{\text{max}}/I_0})$, where $\lambda$ is the wavelength of excitation beam, $NA$ is the numerical aperture of the system, $I_{\text{max}}$ is the peak intensity of the depletion beam and $I_0$ is the threshold intensity of the fluorophore. It shows that the resolution is relative to $I_{\text{max}}$ which determines the depletion area in the excited fluorophores, and can be enhanced further by increasing the ratio of $I_{\text{max}}$ to $I_0$ but below the damage threshold of fluorophores. The most recent STED based on solid immersion lens has the capability of achieving the imaging resolution down to 2.4 nm.[38] Since the invention of STED, many improvements such as multicolor channel,[19] 3D and video rate imaging of living cells, have already been developed subsequently.[40-47]

Another type of superresolution fluorescence techniques adopted broadly is single molecule switching and localization microscopy, which includes photoactivated localization microscopy (PALM) proposed by Betzig et al.[48-50] and stochastic optical reconstruction microscopy (STORM) developed by Zhuang and co-workers.[51-54] These approaches are based on the stochastically photoswitching or photoblinking of individual fluorophores between on and off states. Under the illumination of excitation beam, parts of fluorophores will be emitting or blinking in each piece of time, which facilitates the accurate localization of the fluorophores after the localization algorithm. By implementing thousands of consecutive cycles of stochastic photoswitching, the localization density of the fluorophores with very high spatial resolution can be obtained. Then an image of the fluorophore-labelled specimen is reconstructed by merging the locations of all the individually activated fluorophores. The typical resolution of the PALM and STORM could be down to tens of nanometers.[54-56]

In addition, several other techniques, such as saturated structured illumination microscopy (SSIM),[57,58] superresolution optical fluctuation imaging,[59,60] ground state depletion,[61,62] MINIFLUX,[63] etc., could also realize imaging beyond the diffraction limit. All these technologies are based on fluorescent molecules and have a selective dependence on specimen. It is crucial to label the biological sample appropriately. Moreover, the photobleaching is also an inevitable issue to be considered in the fluorescence-based microscopy.

1.3. Label-Free Far-Field Microscopies

All those superresolution techniques discussed above rely on either critical near-field operation or specimen preprocesses. In recent years, we have witnessed various advances in the development of far-field superresolution label-free imaging technologies.[64-74] A sparsity-based coherent diffractive imaging (CDI) technique[71,74,75] can reconstruct the image of some sparse nano-objects from their multiple spatial spectra measured at the far field. During the measurements, their phase profiles and all the information from evanescent waves are lost. Their retrieval is realized by using compressed sensing algorithms, so that a digital image of these nano-objects can be obtained eventually.[71] In compressed sensing algorithm, the retrieval is an optimization process that minimizes the difference between the measured and optimized spatial spectra. The limitation of such a technique is that it will fail when two slightly different deep-subwavelength objects have the nondistinguishable spatial spectra in measurement. Its applications are also constrained to imaging sparse objects.[71]

Another method for far-field label-free microscopy is a purely optical approach by using a planar diffractive lens (PDL) integrated into confocal scanning microscopy. The focused light spot by the PDL works as an optical noninvasive probe to illuminate the sample at the far field and directly determines the resolution of PDL-based confocal scanning microscope.[72,73] Benefiting from the rapidly developed nanotechnology and powerful optimization algorithms, the planar lenses composed of micro or nanostructures can be customized with many intriguing functionalities such as ultrahigh numerical aperture,[76] large depth of focus (DOF),[76,77] and subdiffraction-limit focusing in the far-field,[78-81] allowing for superresolution imaging. Because the PDLs have large degree of freedom in manipulating the focal fields, rapid development with special designs for various applications is foreseeable in the near future. With compact size, flat surfaces, versatile functionalities and powerful designs, the PDLs are also challenging the traditional optical elements such as objective and camera lenses.

Here, we will mainly focus on the recent advances of the fast-growing planar diffractive lenses. This review begins with a general introduction of various superresolution imaging techniques. The diffraction-based focusing optics will be introduced and taken as the physical origin of PDLs, whose focal spots, depth of focus, and diffraction limit in focusing are subsequently discussed in details. We then proceed to discuss the structures and properties of different PDLs from their design principles, fabrication processes, focusing performances, and optical aberrations. A theoretical treatment on PDL-based scanning focal microscopy and direct wide-field imaging is provided with the existing experimental results, followed by a detailed comparison among all these optical superresolution microscopies. Finally, this review is summarized with a short comment on the current research status, challenges, and possible directions for future applications.

2. Diffraction-Based Focusing Optics

2.1. Mathematical Fundamentals of Diffraction

Diffraction is an intrinsic phenomenon in water, electromagnetic, and acoustic waves during their propagation in free space and media. It can be explained by Huygens–Fresnel principle describing a new wavefront as “a superposition of wavelets originating from an infinity of fictitious ‘secondary’ point sources” in the primary wavefront.[5] Such a principle in electromagnetic wave was mathematically formulated into an integral by Gustav Kirchhoff and then modified as Rayleigh–Sommerfeld diffraction theory, which can be rigorously derived from Maxwell equations.[5] For a monochromatic wave with a given electric field $U(x_0, y_0)$ where $x_0$ and $y_0$ are the spatial coordinates, its...
Equation (2) aims at solving the diffraction of a rotational-symmetry lens such as zone-plate-based lenses. It is valid for any diffraction problem that can be described in a polar or cylindrical coordinate. It should be noted that, to use Equation (2), the incident field $U(r, \phi)$ must be formulated analytically. In other words, Equation (2) has an analytical incident field $U(r, \phi)$ and a pixelate output field $u(\rho, \phi, z)$. It means that Equation (2) cannot be used to describe the diffraction of two cascaded rotational-symmetry lenses, because the pixelate output field of the first lens cannot be taken as the incident field of the second lens.

Equation (4) is able to calculate the diffraction of a pixelate or discrete field modulated by holograms or metasurfaces. Since both incident and output fields are pixelate, Equation (4) is valid for single or cascaded metasurface lens, and a metasurface lens focusing light in multilayered media. When describing the propagation of light through different media, one can change the parameter $n$ in Equation (4) by addressing its corresponding refractive index of medium. Equation (4) is more powerful because it can simulate all the diffraction issues in the area of metasurfaces.

The diffraction theory mentioned above is based on scalar optics without considering the vector properties of light. In fact, the vectorial Rayleigh–Sommerfeld diffraction\cite{81} has the similar form but contains more electric components that can be derived from Maxwell equations. A detailed summary of vectorial Rayleigh–Sommerfeld diffraction has been provided in various coordinates,\cite{81} which is a good guideline for investigating the diffraction of a vector beam such as cylindrical vector beams\cite{82,83} with spatially varying states of polarization.

### 2.2. Diffraction from a Ring Belt and Lens Design

Traditional objective lenses utilize the refraction phenomenon happened at the interface between two media to focus light into a spot by tuning the surface parameters such as their curvatures,\cite{86} as shown in Figure 1a. Fundamentally, the diffractive lenses employ the diffraction of light from a microscale or nanoscale ring (see Figure 1b), which is usually adopted due to the rotational-symmetry existing in most optical systems. Such diffractive lenses are usually made in a thin film (at the order of hundreds of nanometers) with its thickness negligible compared with their transverse dimension of lenses.\cite{72,73,76,78} Therefore, we name them as PDLs with micro-/nanostructures to distinguish from the traditional refraction objective lenses with bulky volume, although the refraction objective lenses also have the diffraction effect in focusing light. In this review, the mentioned PDLs mainly contain zone-plate-based lenses, photon-sieve-based lenses, and metasurface lenses. Different from the refraction-based lenses, these PDLs realize the focusing of light through elaborately tailoring the interference of light diffracting from all the elements arranged in multiple concentric rings by optimizing their locations. Therefore, a single ring is an important diffractive unit in all PDLs and its diffraction behavior directly determines the focusing performance of a PDL.\cite{100}

For a transparent ring with a width of $\Delta r$ and central radius of $r_0$ in Figure 1c, its diffraction field can be described by using Equation (2). It should be noted that the transmitted pattern depends on the incident polarizations when the ring width $\Delta r$...
is smaller than one wavelength. For simplicity, we assume that the transparent ring is taken as a ring source with a uniform azimuthal intensity profile, which is reasonable because it can be realized physically by shining an unpolarized light onto such a ring. In this case, light in the transparent area of ring will be taken as unity in the simulation. To obtain a well-confined focal spot, such a ring source is expected to have a diffracted intensity profile similar to zero-order Bessel function at the target plane, see Figure 1d showing its maximum intensity located on the optical axis (i.e., \( r = 0 \)). In contrary, such rings having bad focusing performance like the one in Figure 1e should not be involved in a PDL.

As a roadmap, Figure 1f illustrates the difference, evaluated by the root-mean-square error (RMSE), between the diffracted intensity and zero-order Bessel function for the rings with various width \( \Delta r \) and central radius \( r_0 \) (where \( r_0 \) is denoted in terms of \( \sin \alpha = r_0 / (r_0^2 + z^2)^{1/2} \) when \( \Delta r \) is small. Intensity profiles (red dots) of light diffracted from belts with \( \Delta r \) small (\( \Delta r = 0.5 \lambda \)) and \( \Delta r = 1.7 \lambda \) widths. The blue solid curves show the square of the zero-order Bessel function of the first kind. \( z = 20 \lambda \). f) Difference between intensity profile at the target plane and its zero-order Bessel function \(|J_0(kr \sin \alpha)|^2\) with the same \( \alpha \). The difference is evaluated by their root-mean-square errors (RMSEs), where only the cases with small RMSEs are shown in color. c–f) Reproduced with permission. [100] Copyright 2014, Wiley-VCH.

Figure 1. Diffraction-based focusing theory. Sketch of optical focusing by a) a refraction-based objective lens and b) a diffraction-based planar lens. c) A diffraction unit of a single transparent belt with the ring width \( \Delta r \), central radius \( r_0 \), maximum convergent angle \( \alpha \) at the target plane \( z \). \( \sin \alpha = r_0 / (r_0^2 + z^2)^{1/2} \) when \( \Delta r \) is small. Intensity profiles (red dots) of light diffracted from belts with d) small (\( \Delta r = 0.5 \lambda \)) and e) large (\( \Delta r = 1.7 \lambda \)) widths. The blue solid curves show the square of the zero-order Bessel function of the first kind. \( z = 20 \lambda \). f) Difference between intensity profile at the target plane and its zero-order Bessel function \(|J_0(kr \sin \alpha)|^2\) with the same \( \alpha \). The difference is evaluated by their root-mean-square errors (RMSEs), where only the cases with small RMSEs are shown in color. c–f) Reproduced with permission. [100] Copyright 2014, Wiley-VCH.

For a PDL, the lateral size of its focal spot can be customized by optimizing the structure (i.e., width and radial position of every
ring) of a PDL to meet special requirements in various applications. Figure 2a gives a detailed summary about 1D (along x) intensity profiles of various focal spots. A traditional Fresnel zone plate (FZP) without further optimization usually focuses light into an Airy spot with the size of \(0.61\lambda/NA\), which is the same with Rayleigh criterion defined in a spherical-lens-based optical imaging system. \([6]\) In addition to the main spot, the Airy spot still has a weak sidelobe whose peak intensity is 1.75% of that of main spot, \([72]\) as shown in Figure 2a. This sidelobe can be further suppressed by a photon-sieve-based lens \([86,87]\) while it will lead to a larger main spot above Rayleigh criterion, which means that the radial dimension of interested field is much smaller than the distance from the ring to on-axis point of the cut plane \(z_0\). In Equation (5), we use these approximations: 1) the item \(R^2 = \sqrt{r_{\text{max}}^2 + z_0^2}\), the integral along \(r\) is ignored due to small \(\Delta r\) and \(R^2 = r_{\text{max}}^2 + \rho^2 + z^0 - 2r_{\text{max}}\rho\cos \phi\). To further simplify Equation (5), we introduce the “near-axis” approximation of \(\rho \ll \sqrt{r_{\text{max}}^2 + z_0^2}\), which means that the radial dimension of interested field is much smaller than the distance from the ring to on-axis point of the cut plane \(z_0\). The focal fields of a lens composed of three \((r_1 = 0.7\lambda, r_2 = 1.9\lambda, \text{and } r_j = 30\lambda)\) belts with their \(\Delta r = 0\). The spot size of \(0.172\) is obtained at the plane of \(z = 10\lambda\), while the depth of focus is \(0.2\lambda\). b) Reproduced with permission. \([100]\) Copyright 2014, Wiley-VCH.

c) Reproduced with permission. \([72]\) Copyright 2017, Wiley-VCH.

generally, optical superfocusing beyond Rayleigh criterion brings more practical benefits but at the cost of the increasing sidelobes (see Figure 2a). To realize it, careful optimizations and elaborate designs of a lens are needed by adjusting the proportion of high and low spatial frequencies involved in the focusing, because light with higher spatial frequencies corresponds to a smaller main spot. The extreme case is that, light with only the maximum spatial frequency can be focused into a hotspot, which is named as “maximum-frequency spot.” To obtain its mathematical formula, Equation (2) is employed under the condition of the ring width \(\Delta r \rightarrow 0\). Thus, the electric field of “maximum-frequency spot” can be expressed as

\[
u_{\text{Max freq.}}(\rho, z_0) = \frac{-ikn}{2\pi} \Delta r \cdot r_{\text{max}} \cdot z_0 \int_{\frac{\rho}{R}}^{\frac{\rho_0}{R}} \frac{e^{ik\rho \rho_0 \rho} \rho_0^2}{R^2} \left( \frac{kn - 1}{R} \right) d\rho
\]

where \(r_{\text{max}}\) is the central radius related to the maximum spatial frequency, \(z_0\) is the longitudinal position of interest, light at the ring is assumed to be uniform and rotation-symmetry, the integral along \(r\) is ignored due to small \(\Delta r\), and \(R^2 = r_{\text{max}}^2 + \rho^2 + z_0^2 - 2r_{\text{max}}\rho\cos \phi\). To further simplify Equation (5), we introduce the “near-axis” approximation of \(\rho \ll \sqrt{r_{\text{max}}^2 + z_0^2}\), which means that the radial dimension of interested field is much smaller than the distance from the ring to on-axis point of the cut plane \(z_0\). In Equation (5), we use these approximations: 1) the denominator \(R^2 = \sqrt{r_{\text{max}}^2 + z_0^2}\); 2) in the numerator \(\exp(inkR)\), the item \(R = \sqrt{r_{\text{max}}^2 + z_0^2} \cdot \sqrt{1 + \frac{\rho^2 - 2r_{\text{max}}\cos \phi}{2(r_{\text{max}}^2 + z_0^2)} + \frac{2}{2(r_{\text{max}}^2 + z_0^2)} \cdot 1 + \frac{\rho^2 - 2r_{\text{max}}\cos \phi}{2(r_{\text{max}}^2 + z_0^2)}} = \sqrt{r_{\text{max}}^2 + z_0^2} \cdot \left(1 - \frac{r_{\text{max}}\cos \phi}{r_{\text{max}}^2 + z_0^2}\right) + 3) the item \(1/R(\ll kn)\) is ignored. Thus, Equation (5) can be simplified as

\[
u_{\text{Max freq.}}(\rho, z_0) = \frac{-ikn}{2\pi} \Delta r \cdot r_{\text{max}} \cdot z_0 \int_{\frac{\rho}{R}}^{\frac{\rho_0}{R}} e^{ik\rho \rho_0 \rho} \rho_0^2 \left( \frac{kn - 1}{R} \right) d\rho
\]

where \(\sin \alpha_{\text{max}} = r_{\text{max}}/(r_{\text{max}}^2 + z_0^2)^{1/2}\), the numerical aperture \(NA = n\sin(\alpha_{\text{max}})\) and we use the equality \(\int_{\frac{\rho}{R}}^{\frac{\rho_0}{R}} e^{ik\rho \rho_0 \rho} \rho_0^2 d\rho = -2\pi J_0(k\rho \rho_0 \rho)\). For the first kind, \(J_0\) is the zero-order Bessel function of the first kind. Equation (6) shows that the electric field of this “maximum-frequency spot” can be mathematically approximated as a zero-order Bessel function under the “near-axis” condition. One can easily find that, the simulated results in Figure 1d verify that Equation (6) holds...
well when the radial position is smaller than 2λ, which is large enough to evaluate the lateral size of spot.

This maximum-frequency hotspot has a main spot with the size of 0.38λ/NA (i.e., the first zero-point of J0) and the peak-intensity ratio of 16.2% between the strongest side lobe and main lobe, which is still acceptable in applications due to the good trade-off between small main-spot size and weak side-lobe intensity.

Furthermore, the main spot can be further reduced so that light at the main-spot region oscillates faster than the maximum-frequency hotspot. This phenomenon is called as superoscillation in mathematics.[112–119] Correspondingly, the item 0.38λ/NA kept in maximum-frequency hotspot is taken as superoscillation criterion, which means that a superoscillatory spot has its size smaller than 0.38λ/NA (see Figure 2a). A detailed discussion about the definition of superoscillation criterion in optical focusing can be found in our work.[100] According to the superoscillation theory, the size of main spot can be infinitely small without any limitation in principle.[80,120–122] However, such a superoscillatory spot has strong sidelobe that its effective field of view in optical imaging is confined within the limited area enclosed by the sidelobe.[123,124] This limitation has to be solved before it can be used in large area imaging. Figure 2b is a universal map describing the relationship between focal size and numerical aperture of a lens with the help of two important boundaries: Rayleigh and superoscillation criterion. The yellow region means that an above-Rayleigh criterion spot has its lateral size larger than 0.61λ/NA. The cyan region covers a subdiffraction spot with its size located between both criteria. The blue region denotes those superoscillation spots having the sizes below superoscillation criterion 0.38λ/NA. In combination with Figure 2a, one can find that the sidelobe of a focal spot increases smoothly as the lateral size of main spot decreases across these three regions.

### 2.4. Depth of Focus and Optical Needle

A lens has another important factor of the DOF that directly influences its performance in imaging and lithography. The DOF discussed here is the longitudinal (or the propagation direction of light, as denoted as z direction here) range around the designed focal plane, where the pursued focal spot is well-maintained without any loss in optical performance such as the lateral spot size. In various applications, the preferred DOFs are different, depending on their special requirements. For example, 3D lithography[125,126] needs a smaller DOF to obtain a volume-suppressed spot for enhancing the axial resolution in fabrication. But, in a scanning confocal microscopy (SCM),[127–131] larger DOF is preferred so that the generated spot can illuminate the samples with a good tolerance to spatial deviation and obliquity.

For a traditional lens system in air or vacuum, its DOF can be described as DOF = λ/(1 − cosαmax) by using the optical analog of uncertainty principle.[6,130] The largest convergent angle αmax between the on-axis ray and relative axis is to the maximum spatial frequency of this system with its NA = sinαmax. For a low-NA (i.e., small αmax) system, its DOF can be approximated as 2λ/sin2(αmax), which is commonly adopted in lens design and literatures.[6,84] Such DOF is also valid for a standard Fresnel zone plate with its ring radii obeying[84]

\[ r_n = \sqrt{n\lambda f + (n\lambda/2)^2} \]  

(7)

where f is the focal length and n is the index of ring, and a metasurface lens with a fixed phase profile of[85,94,113]

\[ \phi (r) = \frac{2\pi}{\lambda} \left( f - \sqrt{r^2 + \alpha^2} \right) \]  

(8)

where f is the focal length and r is the radial coordinate.

In comparison, the DOF of a PDL with optimized structures can be customized at a certain but physically reasonable value by using careful designs. Categorized by the lateral size of focal spot, the DOF of a PDL will be discussed in two scenarios. First, for a PDL with its focal spot larger than the superoscillation criterion 0.38λ/NA, its DOF reported in both theory[81,106,107,134–136] and experiment[2,7,6,79,137,138] can take the value of several tens of wavelengths (even for a high-NA PDL), which is mainly attributed to the powerful optimization technique. Based on constructive interference, such a long DOF can be taken as an optical subwavelength needle, which also exists in a high-NA objective lens with additional phase[106,107] or amplitude[139–141] modulation. In physics, this subwavelength needle is very interesting because it can propagate without any significant divergence and keep its well-confined shape for several tens of Rayleigh lengths[84] (i.e., zR = πw02/λ, where w0 is the lateral size of a needle) that is usually used to evaluate the divergence of a laser beam. For example, an experimentally measured 12λ-long needle with its lateral size of around 0.407λ/θ (see Figure 2c) can propagate over ≈23 Rayleigh lengths. It operates like optical spatial solitons[142,143] and anisotropic media[146,147] to achieve a divergence-free well-confined beam. Such a PDL-based subwavelength needle might provide a new physical insight for the soliton research.

To highlight the significance of long DOF in optical imaging, a supercritical lens[72] has recently been reported to possess a focusing spot slightly larger than superoscillation criterion (0.38λ/NA), and a needle-like focal region with its DOF larger than 2λ/NA2 that distinguishes from traditional objective, FZP and superoscillation lenses. This supercritical lens stands for a nearly perfect balance among the subwavelength focal spot, weak sidelobe, long DOF and large focal length for practical applications. Integrated into a confocal scanning microscopy, this supercritical lens has the overwhelming advantages in improving the imaging resolution and mapping a 3D object over some traditional microscopy technologies, as introduced in details latter.

Second, the DOF of a superoscillation lens, taken as one of the PDLs, is quite limited due to the physical origin of destructive interference occurred in a superoscillation field. Figure 2d shows the cross-section (x–z plane) of a propagating field having a superoscillation spot (with a lateral size of 0.17λ) located around z = 10λ. The on-axis (x = 0 and y = 0) intensity has a valley at the designed cut plane and increases quickly
for a z-direction spatial deviation, accompanying with the fast increment of the lateral size of the spot. This leads to a limited longitudinal region of $\approx 0.2\lambda$ where the superoscillation happens (see Figure 2d). Such a phenomenon has also been observed in experiment, which indicates that the destructive interference in a superoscillation field needs more rigorous conditions than the constructive interference required in a sub-wavelength needle. It means that the rigorous destructive interference cannot hold at an observation plane slightly deviated from designed z-plane because its phase and amplitude profiles from multiple interfered beams also change. To the best of our knowledge, there has been no report about a free-space superoscillation needle with its lateral size smaller than $0.38\lambda/NA$. It should be pointed out that, although the phrase “superoscillatory needle” has been used in some literatures, those generated or proposed needles have the lateral size above the superoscillation criterion and the superoscillatory spots are obtained beyond the needle region. In addition, M. V. Berry, the pioneer in this superoscillation area, has reported that a superoscillation field containing the subwavelength details of an object has the DOF at the $10^{-45}\lambda$ order of magnitude during its propagation.

2.5. Diffraction Limit of a PDL

For an objective lens, its diffraction limit in optical focusing or imaging is widely accepted to be Rayleigh criterion of $0.61\lambda/NA$ (or $0.51\lambda/NA$ where the spot size is evaluated by its full width at half maximum (FWHM)). In fact, such a diffraction limit can be easily broken by combining the objective and some additional diffractive optical elements with careful optimization. Therefore, the Rayleigh criterion has not precisely predicted the real diffraction limit of light, especially for a PDL with optimized structures. The above discussions about the focal spot and DOF suggest that the new diffraction limit should be updated as the superoscillation criterion of $0.38\lambda/NA$. The considerations are in three aspects. First, when the spot is smaller than $0.38\lambda/NA$, its sidelobe gets stronger and can even exceed the main lobe. This sidelobe sets an extremely limited field of view in imaging and is also a big barrier in nanolithography. Second, the short DOF ($<\lambda$ among all reported literatures) of a superoscillation spot makes the optical alignment challenging in practical applications, which results in poor tolerance to the spatial deviation and obliquity of samples. Finally, when the size of the main focusing spot in the superoscillation lens is deep-subwavelength, its intensity is so weak that it could be taken as noise compared with the sidelobe. Therefore, it is reasonable to define the diffraction limit of a PDL as $0.38\lambda/NA$.

3. Planar Diffractive Lenses

In order to distinguish different types of PDLs, we classify them by their micro-/nanofeatures as zone-plate-based lenses, photon-sieve-based lenses, and metasurface lenses. Among them, the representative ones are specifically highlighted in Figure 3, which gives a detailed comparison by addressing their optical performances such as the size of focal spot, ambient medium, DOF, focal length, numerical aperture, operating wavelength, and the functionalities.

3.1. Zone-Plate-Based Lenses

Standard Fresnel zone plates, as a good candidate of zone-plane-based lenses, have been widely investigated and used in many fields. Therefore, here we only review the recent advances in zone-plane-base lenses composed of many concentric transparent and optimized belts. These novel lenses show the superior performances such as tighter focusing spots and broadband operation compared with the standard FZPs. They are categorized by the modulation types into amplitude and phase zone-plane lenses, as discussed below.

The amplitude zone-plane lenses are frequently demonstrated due to its easy fabrication by etching the belts in a thin metal (e.g., gold, silver, chromium, etc.) film. A superoscillatory lens composed of 25 transparent belts with the feature widths of 200 nm (smallest) and 1200 nm (largest) has been experimentally demonstrated with a focal spot of $0.29\lambda$ (FWHM) in oil and a focal length of $16.1\lambda$, see Figure 3a. As one of the most intriguing demonstrations, this high-NA superoscillatory lens excites intensive researches on various zone-plane-based PDLs. A similar binary-amplitude lens was aimed at generating a subdiffraction-limit optical needle for the linearly and circularly polarized light. Then, a microstructured lens with a diameter of 978.4 µm was designed for focusing azimuthally polarized vortex light into a $12\lambda$-long optical subwavelength needle with an impressive focal length of 240µ at the wavelength of 633 nm, see Figure 3b. Due to the feature size of micrometer scale, it can be fabricated by using laser direct writing or standard UV photolithography for mass production, without the requirement of high-cost and time-consuming electron-beam or focused-ion-beam lithography. A 0.98NA microscale lens was then demonstrated as the first prototype of supercritical lens with its focal spot close to $0.38\lambda/NA$ and a long DOF at the violet light, which has the great advantages in noninvasive imaging as shown later. In principle, at different wavelengths such as red, green and blue light, the needle light can also be generated with a slight longitudinal deviation, named as chromatic dispersion. But, the longitudinal extension of these colored needles makes them overlay partially over a certain region where all three colored spots can be obtained at an identical z-cut plane. In addition, a low-NA 3D lens composed of three longitudinally stacked FZPs that have the same focal lengths at three primary colors respectively, was also proposed to correct the chromatic aberrations, see Figure 3c. Although these amplitude lenses have excellent performances, their total efficiencies are limited below 10%.

The binary-phase zone-plated lenses, with the enhanced efficiency of around 40% in theory, have also been reported for tailoring the radially and circularly polarized light, see Figure 3d. Different from the amplitude lenses, these phase lenses fabricated on a transparent dielectric (e.g., SiO₂, Si₃N₄) film have the size-varying grooves, which introduce the required phase difference of $\pi$. Focused by a traditional objective or a PDL with high NA, a radially polarized beam has a strong longitudinally polarized needle with a subwavelength lateral size in the focal region. In ref. [157], the longitudinally polarized (i.e., $E_z$)
needle measured by using a nanofiber-based NSOM is the first experimental evidence since its theoretical prediction in 2009.[107]

The component $E_z$ of electric field cannot be directly measured in an objective-based system due to the polarization filtering effect caused by the magnification of an imaging system.[6,160–162] It has recently been well-solved by some approaches such as $E_z$-enhanced second harmonic generation[161] and Raman spectroscopy,[163] which can be found in more details in ref. [164].

### 3.2. Photon-Sieve Lenses

Photon sieves as proposed in 2001 are the holes penetrated through an opaque metal film.[86] They were initially designed to suppress the sidelobes in the focal spot. Recently, it has been found that the nanoscale circular holes have no effect of polarization dependence that exists in a zone-plate-based lens with purely subwavelength widths,[88,165] which should be distinguished from the superoscillatory lens composed of wavelength-scale rings in ref. [73]. A fractal photon-nanosieve lens,[87] with the nanoholes located within the belt regions of FZPs, was designed to suppress both side lobes and high-order diffraction, see Figure 3e. Due to the low NA, this lens has a diffraction-limited focusing spot. To achieve a smaller spot, more nanoholes are needed at the outmost belts to include light with higher spatial frequencies, which eventually increases the total size of lens and brings a challenging issue in its simulation and design. When the hole is at the scale of nanometers, the rigorous vector theory

<table>
<thead>
<tr>
<th>Lenses Types</th>
<th>Refs.</th>
<th>SEMs</th>
<th>Spot (FWHM)</th>
<th>Ambient Medium</th>
<th>DOF</th>
<th>Focal Length</th>
<th>NA</th>
<th>Wavelength (λ)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone Plates</td>
<td>Rogers et al.[73]</td>
<td>a</td>
<td>0.29λ</td>
<td>Oil</td>
<td>Not Reported</td>
<td>16.1λ</td>
<td>1.34</td>
<td>640nm</td>
<td>Focusing</td>
</tr>
<tr>
<td></td>
<td>Qin et al.[72]</td>
<td>b</td>
<td>0.41λ</td>
<td>Air</td>
<td>12λ</td>
<td>135λ</td>
<td>0.98</td>
<td>405nm</td>
<td>Focusing</td>
</tr>
<tr>
<td></td>
<td>Avayu et al.[159]</td>
<td>c</td>
<td>5.78λ 4.42λ</td>
<td>Air</td>
<td>Not Reported</td>
<td>1818λ(550nm)</td>
<td>0.11</td>
<td>450nm 550nm 650nm</td>
<td>Focusing &amp; Imaging</td>
</tr>
<tr>
<td></td>
<td>Yu et al.[157]</td>
<td>d</td>
<td>0.4λ -0.52λ</td>
<td>Air</td>
<td>5λ</td>
<td>200λ</td>
<td>0.93</td>
<td>633nm</td>
<td>Focusing</td>
</tr>
<tr>
<td>Photon Sieves</td>
<td>Liu et al.[87]</td>
<td>e</td>
<td>1.03λ</td>
<td>Air</td>
<td>4.5λ</td>
<td>17.5λ</td>
<td>0.48</td>
<td>640nm</td>
<td>Focusing</td>
</tr>
<tr>
<td></td>
<td>Huang et al.[88]</td>
<td>f</td>
<td>0.32λ</td>
<td>Air</td>
<td>Not Reported</td>
<td>21.2λ</td>
<td>0.83</td>
<td>633nm</td>
<td>Focusing</td>
</tr>
<tr>
<td>Metasurfaces</td>
<td>Aieta et al.[96]</td>
<td>g</td>
<td>129λ</td>
<td>Air</td>
<td>Not Reported</td>
<td>1936λ</td>
<td>0.015</td>
<td>1550nm</td>
<td>Focusing</td>
</tr>
<tr>
<td></td>
<td>Arbabi et al.[97,99]</td>
<td>h</td>
<td>0.57λ 0.91λ</td>
<td>Air</td>
<td>Not Reported</td>
<td>32.3λ</td>
<td>0.97</td>
<td>1550nm</td>
<td>Focusing</td>
</tr>
<tr>
<td></td>
<td>Khorasaniejad et al.[89]</td>
<td>i</td>
<td>-0.7λ</td>
<td>Air</td>
<td>Not Reported</td>
<td>1363λ</td>
<td>0.8</td>
<td>532nm</td>
<td>Focusing &amp; Imaging</td>
</tr>
<tr>
<td></td>
<td>Chen et al.[92]</td>
<td>j</td>
<td>0.45λ 0.5λ</td>
<td>Oil</td>
<td>Not Reported</td>
<td>308λ 845λ</td>
<td>1.1</td>
<td>405nm 845nm 532nm</td>
<td>Focusing &amp; Imaging</td>
</tr>
</tbody>
</table>

Figure 3. A summary of various PDLs based on different structures. a–c) Amplitude and d) phase zone-plate lenses contain the basic diffraction elements of concentric belts. e,f) Photon-sieve lenses employ the nanoholes drilled in a thin and opaque film. Metasurface lenses use the g,h) gradient and i,j) geometric phase to realize optical focusing. In the item “Function,” all the lenses are categorized by evaluating their capabilities of imaging an object directly, so that the superoscillation and supercritical lenses are labelled as “Focusing” function because they demonstrate the imaging function with the help of confocal scanning microscopy. Images: a) Reproduced with permission.[73] Copyright 2012, Nature Publishing Group. b) Reproduced with permission.[72] Copyright 2017, Wiley-VCH. c) Reproduced under the terms of the CC-BY Creative Commons Attribution 4.0 International License.[156] Copyright 2017, The Authors, published by Nature Publishing Group. d) Reproduced under the terms of the CC-BY Creative Commons Attribution 4.0 International License.[137] Copyright 2016, The Authors, published by Nature Publishing Group. e) Reproduced with permission.[87] Copyright 2017, Wiley-VCH. f) Reproduced with permission.[88] Copyright 2016, The American Association for the Advancement of Science. g) Reproduced with permission.[96] Copyright 2012, American Chemical Society. h) Left: Reproduced with permission.[99] Copyright 2015, Nature Publishing Group. Right: Reproduced under the terms of the CC-BY Creative Commons Attribution 4.0 International License.[97] Copyright 2016, The Author, Published by Nature Publishing Group. i) Reproduced with permission.[92] Copyright 2017, The American Association for the Advancement of Science. j) Reproduced with permission.[93] Copyright 2017, American Chemical Society.
of diffraction is needed to approximate the diffraction fields of a lens. For a large-scale photon-sieve lens, the commonly used simulation method by finite-difference time domain cannot afford due to the high cost in computing sources.

In order to solve the problem, a hybrid approach has been proposed to describe the vector fields of light diffracted from a nanostructure in an analytical way, which opens the door to manipulate light with high accuracy at the nanoscale. Benefitted from this, the photon sieves with optimized locations have been proposed to demonstrate a high-NA focusing lens with a deep-subwavelength focusing spot beyond superoscillation criterion for a linear polarized light, which is realized by precisely manipulating the interference of diffracted light at the focal plane, see Figure 3f. In addition, a hologram with randomly located nanosieves can be realized with good uniformity, ultrabroadband operation and large angle of view.

Such a nanosieve platform suffers from the low transmission at the visible and infrared wavelengths, which leads to a low efficiency for a nanosieve-based optical device. However, the accurate manipulation of transmitted waves makes the nanosieve platform competent at short-wavelength such as X-ray and matter waves in novel microscopies, due to higher transmission efficiency.

3.3. Metasurface Lenses

Metasurfaces are the spatially varying and ultrathin nanostructures that can realize the full manipulation of light including its polarization and required phase, resulting in a low efficiency. For visible-light metasurfaces, the TiO2 film is prepared on a patterned resist by using atomic layer deposition, followed by a carefully controlled reactive ion etching to eradicate the unwanted TiO2 for the final removal of the resist. Single-crystalline Si with a low absorption at the visible and near-infrared wavelengths is also utilized as metasurfaces with the efficiency of 71%. To realize a high-efficiency reflective metasurfaces, one might refer to these nanostructured metasurfaces made on the top of uniform metal and insulator films.

These metasurfaces imparted with a fixed phase profile of \( \phi(t) = 2\pi/\lambda \left[ f - (r^2 + \hat{f})^{1/2} \right] \), as mentioned above, have their focal spots limited by Rayleigh criterion. Such a phase profile enables light passing through metasurfaces to own the same optical path at the designed focus, leading to the constructive interference. In fact, the optimization technique is needed to diminish the size of the focal spot of a metasurface, as the zone-plate and photon-sieve lenses have done. Therefore, it is reasonable to expect a tighter focusing spot beyond Rayleigh criterion by using an optimized metasurfaces considering that all the PDLs such a metasurfaces, a zone-pate or a photon-sieve lens, have the similar optical aberrations, the next section will be fully focused on the aberrations of PDLs to highlight the significance of correcting them for practical applications.

4. Optical Aberrations

In a PDL-based system, the aberrations mainly originate from the off-axis and multilayer illumination. But, a traditional objective lens has another on-axis aberration, named as spherical aberration, which refers to that, the divergent rays coming from an on-axis point source are focused at the different longitudinal (or axial) locations that depend on their divergent angles, due to the imperfection of a spherical lens. The spherical aberration is absent in a single-wavelength PDL because light incident on the inner and outmost regions of a PDL is designed or optimized to have the same focal position.

4.1. Off-Axis Aberrations

Optical imaging by using a single PDL is feasible to map an object onto the designed image plane, with the object and image distances labeled by \( l_1 \) and \( l_2 \) respectively (see Figure 4a). Correspondingly, the structures of a PDL are constructed to guarantee the constructive interference, that is, all rays coming from an on-axis object have the same optical paths at the
designed image positions. In this configuration, the relative imaging FZPs have their radii obey
\[ \sqrt{r_1^2 + r_2^2} + \sqrt{r_1^2 + r_3^2} - l_1 - l_2 = \frac{n}{2} \lambda \]  
(9)

Similarly, the phase profile of a metalens is modified as
\[ \phi(r) = \frac{2\pi}{\lambda} \left( l_1 + l_2 - \sqrt{l_1^2 + r^2} - \sqrt{l_2^2 + r^2} \right) \]  
(10)

We have to emphasize that Equations (9) and (10) are valid for the on-axis objects and have the aberrations for the off-axis objects.

To show the off-axis aberrations in an imaging PDL, we assume the object \( P \) is located at \( (\xi, \eta) \) without the loss of generality (see Figure 4a), because any off-axis object is composed of infinite points. The off-axis aberrations in a PDL can be evaluated by using the optical path difference between the central (blue dashed line) and marginal (blue solid lines) rays emitting from an off-axis point source \( P \) to its ideal image point \( Q \). After a brief mathematical derivation, the optical path difference can be expressed as
\[ \Delta l = \frac{l_1 + l_2}{\cos \theta} - \sqrt{l_1^2 + (R - l_1 \tan \theta)^2} - \sqrt{l_2^2 + (R + l_2 \tan \theta)^2} \]  
(11)

where \( l_1 \) and \( l_2 \) are the object and image distances from the PDL, \( R \) is the radius of PDL, \( \theta \) is the angle between the optical axis and the central ray. Any off-axis aberration in such a PDL can be predicted by using Equation (11).

Taking a metalens for example by using Equations (3) and (4), we show the simulated intensity profiles at the image plane when the off-axis value of the point object is different. For a low NA case in Figure 4b, this metalens-based system has the parameters \( R = 100\lambda \), \( l_1 = 200\lambda \), and \( l_2 = 300\lambda \), resulting in its NA = \( \sin \tan^{-1}(R/l_1) = 0.447 \). Figure 4b shows that the intensity profile of the image gets the increasing distortion when the off-axis distance enlarges. But, the distortion for low-NA case is still acceptable compared with the 0.9NA case as shown in Figure 4c, which...
has the parameters \( R = 100\lambda, l_1 = 50\lambda, \) and \( l_2 = 75\lambda. \) One can find that the severe distortion happens when the deviated distance is larger than two wavelengths for a high-NA case. Correcting the off-axis aberrations of a high-NA metalens more challenging as well as significant because the high-NA lens is more powerful in focusing light into a hotspot and resolving tiny objects.

Beyond the off-axis case shown in Figure 4a, the most frequently used FZPs and metalenses are structured with Equations (7) and (8) respectively under the plane-wave illumination, where the optical path difference to evaluate the off-axis aberrations can also be obtained in Equation (11) by setting \( l_i \rightarrow \infty. \) Thus, we have the optical path difference

\[
\Delta l = \frac{l}{\cos \theta} - R \sin \theta - \sqrt{l_1^2 + (R - l_1 \tan \theta)^2},
\]

which has been used to investigate the optical abberation in a low-NA FZP under the assumption of small \( \theta. \)\(^{[154]} \) The intensity profiles of a metalens under the tilting illumination with the angles of \( 1^\circ \) and \( 2.5^\circ \) have been reported theoretically and experimentally.\(^{[97]} \) To compare with the off-axis aberrations between a metalens and FZPs, we sketch the optical system of FZP under the tilting illumination with an angle of \( \theta \) in Figure 4d, which also shows the simulated intensity profiles at \( \theta = 0^\circ, 1^\circ, \) and \( 2^\circ. \) One can find that both FZPs and metalens have the similar off-axis aberrations.

Efforts have been made to correct the off-axis aberrations in a low-NA system. A metasurface doublet composed of a phase-correction metalens and a focusing metalens (see Figure 4e) has been proposed to suppress the off-axis aberration at 850 nm wavelength and show a well-behaved focusing spot even at a tilting angle of \( 30^\circ \)\(^{[97]} \) (see Figure 4f). Both metalenses with gradient phase realized by size-varying Si nanopillars are fabricated on the top and bottom surfaces of a quartz substrate with good alignment. The correction and focusing phase profiles, as denoted by lens 1 and lens 2 in the insets of Figure 4e, are provided with a 3D view. The correction lens 1 has a smaller diameter than focusing lens 2, so that only a limited part of focusing lens 2 is illuminated as shown by the propagation of light through the metasurface doublet in Figure 4e. The propagation is simulated by using the numerical Rayleigh–Sommerfeld formulas in Equations (3) and (4). Although it brings some problems such as decreased efficiency, bad integration, challenging alignment between metalenses, low resolving power in imaging, such a doublet approach confirms the feasibility of correcting optical aberrations by using multiple metalenses.

Another role of correction lens 1 is to work as an aperture, which alleviates the difficulty in suppressing the off-axis aberration and solves the low efficiency problem of the whole doublet. It means that, if the correction lens 1 has the same size as the focusing lens 2, the aberration might be unable to be well-corrected. Due to the small size of correction lens 1, the efficient NA of this doublet is determined by the diameter of lens 1 (not the focusing lens 2) and the focal length, resulting in a larger focal spot. It implies that the off-axis aberration of metalens is corrected at the cost of low resolution in focusing and imaging. Therefore, it might be impossible to pursue a PDL system containing the superresolving power (i.e., below Rayleigh criterion) and having its off-axis aberrations corrected simultaneously, as pointed out in ref. \(^{[92]} \).

### 4.2. Chromatic Aberrations

In optics, chromatic aberrations usually refer to the wavelength-dependent focal shifts of a lens, as sketched in Figure 5a. To show it clearly in a PDL, we provide the simulated focal spots of a 0.6NA metalens with geometric phase under the illumination of the red (620 nm), green (550 nm), and blue (480 nm) light. Obeying Equation (8), the phase profile designed at the green wavelength is directly applied to the red and blue wavelengths because the geometric phase is dispersion-free.

For a diffractive optical element, the chromatic aberrations mainly come from material, resonance, and propagation dispersions.\(^{[166]} \) Material dispersion\(^{[190]} \) that is, the refractive index of a material is wavelength-dependent, exists in most diffractive optical elements. Resonance dispersion, occurring in photon-sieve lenses and metalenses, refers to the fact that the modulated amplitude or phase obtained through the interaction between optical modes and a nanostructure is relative to wavelength, structure geometry, and refractive index of materials. Both material and resonance dispersions are usually called as a joint name of modulation dispersion because the material dispersion is considered in resonating processes of optical modes. Propagation dispersion means that the phase or amplitude of multiwavelength light have the wavelength-dependent values even after propagating over the same distance, which occurs in all optical systems\(^{[192,191]} \) and cannot be eliminated in principle. Up to now, the chromatic aberrations can be suppressed by three approaches: dispersion compensation, spatial multiplexing of monochromatic devices, and multicycle phase with different dispersion.

The method of dispersion compensation employs the elaborate tuning of the electromagnetic resonances in metasurfaces so that the total dispersion (i.e., a summation of modulation and propagation dispersions) is zero.\(^{[90]} \) Such a concept was demonstrated experimentally by using a beam deflector made of silicon metasurfaces that can tune light at the near-infrared wavelengths of 1300, 1550, and 1800 nm into the same refracting angle.\(^{[90]} \) It behaves well only at the designed wavelengths but fails at other wavelengths. The propagation dispersion depends on the spatial distance between the device and the target plane, which changes in a case-by-case (such as Fresnel or Fraunhofer region) manner. The method is valid only when the modulation dispersion is flexible enough in the tuning range to compensate the propagation dispersion.

Reorganizing multiple monochromatic elements spatially into an achromatic device provides another route for solving the dispersion problem. Three ultrathin FZPs\(^{[154]} \) operating at the red (650 nm), green (550 nm), and blue (480 nm) wavelengths, respectively, are located along the propagation direction of light and can focus light with three colors at the design focal plane, as shown in Figure 5c,d. Benefitted from this, the integrated FZPs can image a three-color object onto a camera without the significant chromatic aberration, as depicted in the inset of Figure 5d. Similarly, an alternative approach by transverse multiplexing of elements for three wavelengths can also focus multicolor light at the same plane.\(^{[92]} \) Figure 5e shows the sketch of such a transversely multiplexed device with the scanning electron microscopy (SEM) images and its measured color spots at the focal plane.
The multicycle phase method provides the phase modulation from 0 to $N\pi$ (the integral $N > 1$) for increasing the flexibility in phase selection among the wavelengths of interest. Such a principle has been used to manipulate the multiwavelength light by using traditional diffractive optical elements,\[193,194\] which employ the accumulated phase of light by going through the height-varying (i.e., longitudinal multiplexing of structures) dielectric reliefs. For the metalens case in Figure 5f, the reflection phase can be tuned precisely with a sufficiently large range by transverse-geometry-varying (i.e., transverse multiplexing of structures) silicon nanostructures.\[195\] For example, for all the nanostructures, three (i.e., $g_1$, $g_2$, and $g_3$) of them generate the phase $\varphi_0$, $\varphi_0 + 2\pi$, and $\varphi_0 + 4\pi$ respectively at the wavelength $\lambda_1$, but have the different phase modulations of $\varphi_1$, $\varphi_2$, and $\varphi_3$ at $\lambda_2$ due to the wavelength-dependent phase derivations, as shown in Figure 5f. It means that one can choose three structures to suppress the dispersion existing at $\lambda_2$ meanwhile maintain all the optical performances at $\lambda_1$ without any change. Based on this, a dispersion-free reflective lens has been experimentally demonstrated at the near-infrared wavelengths from 1450 to 1590 nm. Due to the great degree of freedom, it has been used to control the chromatic dispersion arbitrarily from negative, zero and positive dispersions. If such a mechanism is applied at the visible wavelengths, one might realize various multiwavelength applications such as high-efficiency full-color holography, display, and projector.
These efforts in fighting against chromatic aberrations are quite exciting, but most of them work only for several discrete wavelengths. More solutions are still needed to correct the dispersion over a continuous and wide spectrum, and in a high-NA lens, which are more challenging. To elaborate it, we show that a detailed focal shifts of a metalens (its parameters are the same as those in Figure 4b) enlarge with the increment of the NA of metalens, as shown in Figure 5g. We have to point out that the focal spots are well-maintained when \( NA < 0.8 \). However, the spherical aberrations arise at the un-designed (such as red and blue light in Figure 5h) wavelengths for \( NA > 0.8 \), which indicates that a longitudinally extended needle-like field (not a pursued well-confined spot) is generated in the focal region at the red and blue wavelengths, see Figure 5h. It means that both chromatic and spherical aberrations exist simultaneously in a high-NA lens, which has been observed experimentally in a 0.98NA silicon-nitride metalens.\(^{[196]}\) Due to the existence of both aberrations, we predict that a metalens (phase profiles obeying Equation (8)) with its NA larger than 0.8 can operate efficiently only at the designed wavelength if no aberration is corrected.

5. Applications

A PDL can focus light into a hotspot, which can directly be applied to, e.g., optical imaging and lithography. The section is assigned to discuss the applications of PDLs.

5.1. Nanoimaging

Most of the superfocusing zone-plate and photon-sieve lenses can focus light into a designed spot but could not directly image the objects due to their severe aberrations. For the imaging purpose, such PDLs can be used as the illuminating light source in a SCM setup, and is therefore named as scanning imaging. Conventional FZPs obeying Equation (7) or (9) and metalenses obeying Equation (8) or (10) can be used to image some objects directly. The detailed introduction of both imaging methods is provided below.

5.1.1. Scanning Imaging

First, we introduce the imaging theory of PDL-based SCM by following Sheppard’s method.\(^{[129]}\) Figure 6a depicts the optical

![Figure 6](image_url)

**Figure 6.** PDL-based scanning confocal microscopy (SCM). a) Working principle. b) Simulated resolution of this PDL-based SCM. In the simulations, the spot generated by PDL is assumed to be the ‘maximum-frequency’ spot (i.e., \( J_0(kr) \)), whose intensity profile is provided in the insert. \( R \) and \( M \) are the radius of the aperture and the magnification of the collection objective lens, respectively. c) Imaging results (right) of some random point objects (left) by using a superoscillatory lens (SOL) based SCM. d) The experimental (middle) and simulated (right) images of random point objects (left) by using a supercritical lens (SCL) SCM. e) Complex curved objects. f) 3D objects. The imaging results by traditional transmission microscopy and SCL-based SCM are displayed at the upper and lower lines of right panel, respectively. c) Reproduced with permission.\(^{[73]}\) Copyright 2012, Nature Publishing Group. d,e) Reproduced with permission.\(^{[72]}\) Copyright 2017, Wiley-VCH.
system of this PDL-based SCM. A PDL serves as the focusing lens and generates a well-confined hotspot at the object plane of an objective lens. This hotspot is used to illuminate the nano-objects that are located at the confocal plane of the PDL and the objective lens. With a magnification of $M$, the objective lens functions as a collector of the scattered light from the nano-objects. A circular aperture (or pinhole) with radius $R$ is located at the imaging plane of the objective lens. In theory, the detector is assumed to be located immediately after the aperture, which is reasonable because the detector is used to record only the photon counts of the scattered light as the pixel value of a scanning image. In this optical system, both the PDL and objective lens are well aligned and the objects can move only in the confocal plane with lateral coordinates $(x_0, y_0)$ with the help of a 3D stage. When the objects are scanned in the confocal plane, the detector records the intensity collected by the objective lens for every scanning position of objects. The correlation between the objects’ scanning positions and the detected signals allows for reconstructing the images of the objects with the aid of computer processing software. Assuming that the electric field at the object plane generated by the PDL is $h_2(x_0, y_0)$, the complex amplitude function of the objects at the object plane is $t_0(x_0, y_0)$ and the impulse function of the collecting lens is $h_3(x_0, y_0, x_0, y_0)$, we can derive the amplitude function at the $x_0y_0$ imaging plane:

$$U(x, y; x_0, y_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(x_0, y_0) \cdot t_0(x_0 - x, y_0 - y) \cdot h_2(x_0, y_0; x_0, y_0) \, dx_0 \, dy_0$$

(13)

where $x_0$ and $y_0$ denote the lateral deviation from the original point when the objects are scanned in the confocal plane. In Equation (13), the electric field at the imaging plane of the objective lens depends on the deviation of objects in the con-

focal plane. The aperture permits us to select the partial intensity at the imaging plane as the detected signal. Assuming that the pupil function of the aperture is $P(x_0, y_0)$, we can obtain the reconstructed images of the objects

$$I(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x_0, y_0) \cdot U(x, y; x_0, y_0) \, dx_0 \, dy_0$$

(14)

Equation (14) provides the final formation of the reconstructed images by using such a PDL-based SCM. For the objective lens (with an effective radius $D_0$), its impulse response $h_3(x_0, y_0; x_0, y_0)$ can be approximated as:

$$h_3(x_0, y_0; x_0, y_0) = \frac{D^2}{\lambda d_0 d_i} \frac{J_1(t)}{t}$$

(15)

where $d_0$ and $d_i$ are the object and image distances as depicted in Figure 6a, $\lambda$ is the operating wavelength, $J_1(t)$ is the first-order Bessel function of the first kind and the parameter $t$ is given by

$$t = k \cdot NA \cdot \sqrt{(x_0/M - x_0)^2 + (y_0/M - y_0)^2}$$

(16)

where NA and $M$ are the numerical aperture and magnification, respectively, of the collecting objective lens, the wave number $k = 2\pi/\lambda$.

In order to evaluate the imaging resolution of PDL-based SCM, we adopt its point spread function (PSF) by setting the object to be an infinitesimal point with its amplitude function given by $t_0(x_0, y_0) = \delta(x_0, y_0)$. In this case, Equation (14) can be simplified as

$$I(x, y) = \left| h_3(x, y) \right|^2 \left[ P(x_0, y_0) \otimes h_2(x_0, y_0; x_0, y_0) \right]$$

(17)

where $\otimes$ denotes the convolution operation of $A$ and $B$. Equation (17) yields the PSF pattern of the PDL-based SCM. The FWHM of its PSF is a good measurement of the imaging resolution. To show its performance in resolving power, a maximum-frequency hotspot (see the insert in Figure 6b) with its electric field of $J_0(kr)$, where $r^2 = x_0^2 + y_0^2$, is used as the illuminating source in our simulations. By using the magnification of $M$ of the objective lens and the radius $R$ of the pinhole aperture, we can calculate its PSF and obtain its resolution (in terms of FWHM) for different cases.

Figure 6b presents its corresponding resolution as we vary the $NA_{obj}$ of the objective lenses from 0.6 to 0.9. We observe that the resolution maintains its value around 0.36$\lambda$ without much dependence on the $NA_{obj}$ of the objective lens and the relative aperture $R/M$. As seen in Figure 6b, there is some improvement in resolution by decreasing the relative aperture for $R/M < \lambda$ while such improvement is quite limited to be $<0.05\lambda$. It means that the resolution of a PDL-based SCM is limited to be the size of illuminated spot generated by this PDL. This result completely differs from the conventional SCM in which the resolution depends strongly on the $NA_{obj}$ of the objective lens. Therefore, without the loss of the imaging resolution, one might have huge flexibilities in choosing the objective lens when the PDL-based SCM is employed.

Considering that the illuminating spot used in the simulations of Figure 6b is the maximum-frequency spot, we predict that the imaging resolution (in terms of FWHM) of a PDL-based SCM is limited to be 0.36$\lambda$/NA, where NA is the numerical aperture of the PDL. If the immersion liquid with a refractive index of 1.5 is used in the PDL-based SCM, the imaging resolution will be $\approx 100$ nm at the visible wavelength of $\lambda = 400$ nm.

Next, we show some imaging results by using such PDL-based SCM. Figure 6c shows the nanohole objects and their images in an oil-immersion superoscillation lens SCM, which can resolve two 210 nm diameter holes with a gap of $\approx \lambda/6$ in oil.[73] The high sidelobe in its superoscillatory spot leads to a high background in the final image, although the sidelobe has been suppressed after being mapped by an objective lens in a standard microscopy, as shown in ref.[73]. Among all optical properties of a superoscillatory lens, the sidelobe might be the biggest challenge to obtain high-quality imaging and therefore is taken as one of the limitations for superoscillatory technique.

To avoid these drawbacks in superoscillatory lenses, we have proposed a supercritical lens that can create a $12\lambda$-length optical needle with its lateral size of 0.407$\mu$m.[72] Integrated into an SCM, such a supercritical lens has the imaging results of point objects shown in Figure 6d. It can resolve two 165 nm diameter holes with a 65 nm gap in air. A clear and background-free image is obtained experimentally and agrees well with the...
the lenses have the fixed magnification
object distance imaging metalens with its radius 1 and R
Figure in various because this phase depends on the object and image distance 
NA also holds for a metalens. In order to investigate the imaging properties of traditional FZPs can be found in many literatures.
The imaging equation of a traditional thin lens with a low NA is only valid under the paraxial approximation,[5] which also holds for a metalens. In order to investigate the imaging performance of a metalens with different NAs, the metalens is imparted with the phase profile defined in Equation (10) because this phase depends on the object and image distance in various NAs (even if the imaging equation is not valid).

**Figure 7a** shows the sketch of an optical system based on an imaging metalens with its radius R, object distance l₁ and image distance l₂. The NA of this metalens can be evaluated by 

\[ \sin \alpha = \sin(\tan^{-1}(R/l_2)) \]

where \( \alpha \) is the angle between optical axis and the maximum divergent rays incident on metalens. Different from the optical behavior of an off-axis point in previous Section 4.1, we here employ a 2D object composed of a \( 10 \times 10 \) array of transparent \( 2\lambda \times 2\lambda \) squares with the smallest gap of \( 2\lambda \), see Figure 7b. In our simulations by using Equation (3), the metalens has the parameters \( R = 100\lambda, l_2/l_1 = 1.5 \), and the varying NAs by scanning \( l_1 \) from 500\( \lambda \) (NA = 0.196), 400\( \lambda \) (NA = 0.243), 300\( \lambda \) (NA = 0.316), 200\( \lambda \) (NA = 0.447), 100\( \lambda \) (NA = 0.707), 75\( \lambda \) (NA = 0.8) to 50\( \lambda \) (NA = 0.894). In addition, the lenses have the fixed magnification \( M = l_2/l_1 = 1.5 \), allowing us to exclude the aberrations caused by different magnifications.

Figure 4b shows the simulated imaging results at the image plane. For the metalense with small NA such as \( l_1 = 500\lambda \), the fine details of object cannot be resolved due to the limited resolution. The resolution is enhanced with the increment of NA so that the object can be mapped correctly with the improved imaging quality. As a consequence, a clear image is obtained by using the metalens with \( l_1 = 200\lambda \). However, when NA increases further, the optical aberrations become significant as predicted in Section 4.1, leaving the pillow-shape images with pincushion distortions (see the cases with \( l_1 = 100\lambda, 75\lambda, \) and \( 50\lambda \)). These simulations confirm that the optical aberration is more significant in a high-NA lens than a low-NA one. Therefore, the reported high-NA metalens is usually used to image the objects directly at the large objective distance, which implies that the efficient NA employed in constructing the image is much smaller than the NA carried by the metalens. Thus, the demonstrated images have the large-scale details but little aberrations, as shown in Figure 7c,

d, and e. For imaging resolution, a metalens with an NA of 0.8 has been reported to enable resolving a 450 nm gap between two holes, see Figure 7f. Considering that its focusing spot has the measured FWHM of 375 nm, the imaging resolution of this metalens is shrunk significantly. The measured modulation transfer function shows that light with its spatial frequency larger than 1000 lines mm\(^{-1}\) cannot pass through this metalens, which might give the reason for the decreased resolution in imaging. Fundamentally, dielectric nanostructures in metalens have the low transmission for the incident light with a large tilting angle, which is relative to the large divergent-angle rays (i.e., high spatial frequencies) emitting from an object (see Figure 7a). During the image formation process in a metalens, the lack or decrement of high-spatial-frequency light will definitely lead to the loss of fine details in objects, leading to a decreased resolution in their images.[5] Therefore, limited by the low transmission of dielectric nanostructures at the large incidence angle, the metalens working in the direct-imaging mode might be difficult to achieve the resolution beyond Rayleigh criterion.

As a result, the most recent oil-immersion metalenses have turned to operate in the scanning mode (as shown in Section 5.1.1.) for nanoimaging.[94] Figure 7g shows the image of a 200 nm width nanograting with a gap of 200 nm obtained at the wavelength of 532 nm, showing improved imaging resolution compared with the direct-imaging case in Figure 7f. Although the scanning mode has lower speed than the direct imaging mode, its advantages in high imaging resolution, small aberration and large object imaging will make the scanning mode more practical in many applications.

### 5.1.3. A Detailed Comparison among Various Optical Superresolution Microscopes

In order to make a guideline for potential users, we list the representative microscopes that can resolve the nanoscale objects beyond the diffraction limit of Rayleigh criterion, as shown in Table 1. For the item of resolution, we use the value of the smallest gaps and center-to-center distances between two objects that have been experimentally resolved by these microscopies. The displayed gap and center-to-center data are exported directly from the relative literatures or derived from the demonstrated images. In this table, the superoscillation- and supercritical-lens-based SCMs with high resolving power are included while all the metalens-based microscopies are excluded due to their diffraction-limited resolution in imaging.

### 5.2. Nanolithography

Nanofocusing of light has another important application in nanolithography, which requires the hotspot with small lateral
size and weak sidelobes. The superoscillatory spot is not applicable in lithography due to the existence of the strong halo. In fact, the zone-plate\cite{109} and photon-sieve-based\cite{110} lenses have been employed for nanolithography in X-rays, extreme ultraviolet and visible light. For example, a 0.9 NA photon-sieve lens operating at the 405 nm wavelength has been reported and demonstrated for fabricating a 400 nm period grating.\cite{108}

The structures contained in photon-sieve and zone-plate lenses are based on the standard parameters defined in Equation (7), resulting in a diffraction-limited focusing spot and the fabricated structures at the scale of wavelength. It is expected that the feature size of fabricated samples by an optimized PDL should be scaled down to its optical limit of $0.38\lambda/NA$.

6. Conclusions and Outlook

In the past five years, exciting achievements about focusing light into a hotspot by using optimized micro/nanostructures in a planar surface have been obtained through extensive researches including diffraction-based focusing optics, new diffraction limit of light, superoscillation phenomenon, optimization techniques in optical simulation, and developing novel optical materials such as metamaterials, metasurfaces, photon sieves, and zone plates. These PDLs with compact volume, light weight, planarity, and superior optical performances are challenging traditional objective lenses. They hold great potentials in reshaping the current microscopy and imaging techniques.
industries to meet the requirements on miniaturized optical elements for portable devices, IT, and automation. However, some challenges should be overcome before their final industrial applications.

First, optical efficiency of a PDL is limited in focusing and imaging. The zone-plate and photon-sieve lenses with binary modulation have the theoretical limitation of below 40% in efficiency. Although dielectric metalenses can improve the efficiency up to 80–90%, the low transmission of light with high spatial frequencies makes it difficult to enhance the imaging resolution for a higher NA metalens. Even in focusing, light incident on the outmost of a high-NA metalens is seldom contributed to the focal spot because the large pixel pitches in a metalens lead to the lower diffraction efficiency from incident light to focal spot, compared with a sufficiently sampling metalens as shown in Figure 7h. Second, the problems associated with high-quality imaging still exist in both operating modes. For example, optical aberrations of a metalens operating in a wide-field mode need to be corrected in premium applications such as nanoimaging. The imaging speed of a PDL-based scanning microscopy might need further improvements to obtain a large-scale image with fine details. Third, dielectric materials with high refractive indices and low absorption at short wavelengths such as visible and ultraviolet spectra are lacked for realizing high-efficiency metasurfaces. Furthermore, the robustness of the micro-/nanostructures in the PDLs in real-world application needs to be addressed. Any solution must be considered holistically to maintain the optical performance of the PDLs. Finally, the process for large volume mass production of the large-area PDLs with tense nanostructures has to be built. Also, a huge desire and challenge ahead is to realize the reconfigurable or programmable planar optical lenses and devices for applications in dynamic imaging and display, although some preliminary active nanodevices have been demonstrated by using phase-change materials,[199,200] gate-tunable conducting oxide,[201] electrically controlled plasmonics[202] or stretchable substrate.[203]

Future work might be focused on both fundamental physics and possible solutions confronting with those challenges for industrial applications. The underlying physics addressing the reason why the lateral size of an optical needle cannot be below 0.38λ/NA (or 0.358λ/NA by FWHM) needs more mathematical proofs, which are of great significance because this might reveal some new physics hidden behind the fact. It is still an open question what is the ultimate limit in optical focusing without significant sidelobes, which is just like that the diffraction limit defined by the Rayleigh criterion is now updated by the superoscillation criterion 0.38λ/NA. For enhancing the efficiency of zone-plate or photon-sieve lenses, the thinner film such as 2D materials[204–206] might be a possible solution although the working spectrum is limited. By referring to the traditional objective lenses, multiple or cascaded metasurface-based singlet might provide a viable way for correcting or reducing optical aberrations to meet the practical requirements in imaging and machining. The mapping of the horizontal details of a 3D object realized by the supercritical-lens-based scanning confocal microscopy might open the opportunities to reconstruct a full-space fine detail of an arbitrarily object in a purely optical way, without the need of labeling technique used in bio-imaging. Such a new capability in a microscopy might be added into one of the key features in developing next-generation laser scanning confocal microscopy. Moreover, the relative theory and technology development are required to realize the superresolution focusing through or in a diffusive medium such as living cells and tissues,[207–211] and multilayered media such as recording materials in high-density data storage.[148] Therefore, the studies in this field will grow further and the solutions in various

Table 1. A summary of optical supersolution microscopies.

<table>
<thead>
<tr>
<th>Microscopies</th>
<th>Near-field microscopies</th>
<th>Far-field fluorescence microscopes</th>
<th>Far-field label-free microscopies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
<td>NOSM</td>
<td>PALM/STORM</td>
<td>SIM/SSIM</td>
</tr>
<tr>
<td>Working principle</td>
<td>Evanescent wave</td>
<td>Photo activation</td>
<td>Structured Light</td>
</tr>
<tr>
<td>Working mode</td>
<td>Scanning</td>
<td>Wide field</td>
<td>Wide field</td>
</tr>
<tr>
<td>Resolution</td>
<td>Scanning</td>
<td>Wide field</td>
<td>Wide field</td>
</tr>
<tr>
<td>Gap</td>
<td>&lt;10 nm</td>
<td>&lt;20 nm</td>
<td>&gt;100 μm</td>
</tr>
<tr>
<td>Center-to-center</td>
<td>&lt;75 nm</td>
<td>&lt;45 nm</td>
<td>&gt;100 μm</td>
</tr>
<tr>
<td>Working distance</td>
<td>&lt;50 nm</td>
<td>&lt;&lt;50 nm</td>
<td>&gt;100 μm</td>
</tr>
<tr>
<td>Wavelength (λ)</td>
<td>Full spectrum</td>
<td>UV–IR</td>
<td>Full spectrum</td>
</tr>
<tr>
<td>Sample features</td>
<td>Unlimited</td>
<td>Visible</td>
<td>Full spectrum</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Very low</td>
<td>Modulate</td>
<td>Transmission-contrasted</td>
</tr>
<tr>
<td>Pre-preparation</td>
<td>No</td>
<td>Yes</td>
<td>Low</td>
</tr>
<tr>
<td>Invasiveness</td>
<td>Mechanical damage</td>
<td>Fluorophore contamination and photo damaging</td>
<td>Noninvasive</td>
</tr>
<tr>
<td>References</td>
<td>[11,12]</td>
<td>[16,18]</td>
<td>[23,24]</td>
</tr>
</tbody>
</table>

© 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
applications will expand continually in a quick way in the near future.

Acknowledgements

B.L. is thankful to DSI core fund and A*STAR Science and Engineering Research Council Pharos (grant 1527000025) for support. B.L. also acknowledges support by the Russian Ministry of Education and Science (#14.W03.31.0008). The work was financially supported by the Institute of Materials Research and Engineering and the Agency for Science, Technology and Research (A*STAR) under Grant No. 152 148 0031. K.H. thanks the support from CAS Pioneer Hundred Talent Plan. F.Q. acknowledges the support from the National Natural Science Foundation of China (Grant No. 61705085) and the Guangdong Innovative and Entrepreneurial Research Team Program (Grant No. 2016ZT06D081).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

flat optics, metasurfaces, subdiffraction limit, superresolution imaging, zone plates

Received: August 11, 2017
Revised: December 17, 2017
Published online: