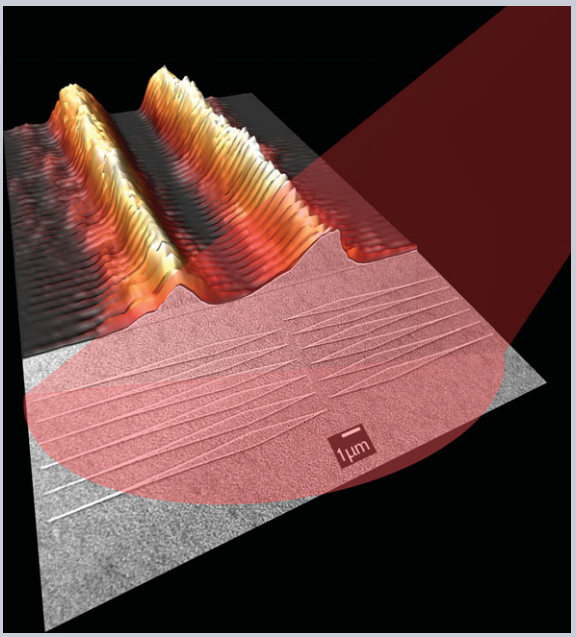
LASER
& PHOTONICS
REVIEWS

Laser Photonics Rev. **10,** No. 3, 360–381 (2016) / **DOI** 10.1002/lpor.201500242

REVIEW
ARTICLE

Abstract Surface-plasmon-polariton waves are two-dimensional electromagnetic surface waves that propagate at the interface between a metal and a dielectric. These waves exhibit unusual and attractive properties, such as high spatial confinement and enhancement of the optical field, and are widely used in a variety of applications, such as sensing and subwavelength optics. The ability to precisely control the spatial and spectral properties of the surface-plasmon wave is required in order to support the growing interest in both research and applications of plasmonic waves, and to bring it to the next level. Here, we review the challenges and methods for shaping the wavefront and spectrum of plasmonic waves. In particular, we present the recent advances in plasmonic spatial and spectral shaping, which are based on the realization of plasmonic holograms for the optical nearfield.



Surface-plasmon wavefront and spectral shaping by near-field holography

Itai Epstein*, Yuval Tsur, and Ady Arie*

1. Introduction

Surface-plasmon-polaritons (SPPs) are surface electromagnetic waves that are coupled to electron waves, and propagate at the surface between a dielectric and a metallic medium [1]. These surface waves exhibit unusual properties, which are usually absent in the case of regular free-space electromagnetic waves, owing to the confinement of the wave to the surface and the enhancement of the optical field. They have been widely used for both fundamental discoveries, such as extraordinary transmission [2], and manipulating Snell's and Bragg's fundamental laws [3,4], and for nanoscale technologies, such as plasmonic circuitry [5], focusing to subwavelength spots [6], sensing [7] and more.

The ability to control and guide plasmonic waves is a crucial step for opening up new possibilities in nanophotonics. Controlling the wavefront properties of surface-plasmon waves, although challenging, is a key feature in enabling a variety of applications. Specifically, in recent years, significant efforts to control the wave properties of plasmons have been performed in order to generate and detect unique plasmonic beams. Among these are "self-accelerating", "diffraction-free" and self-similar plasmonic beams [8–16]. Moreover, the challenging task of controlling the spectral properties of surface-plasmons can open exciting opportunities for broadband coupling, sensing and on-chip manipulation of ultrashort femtosecond pulses.

The pioneering realizations of spatial and spectral shaping of surface-plasmons were mostly specified and relying on ad hoc methods, as they were commonly aimed at a specific application. However, these realizations acted as a major driver for finding more generalized approaches for controlling the plasmon wave properties.

In this manuscript, we will review the different methods applied to surface-plasmon waves in order to perform spatial and spectral shaping of plasmonic beams. Specifically, we will focus on utilizing holographic methods for this purpose. We will introduce the basic concepts of holography and how they can be applied to plasmonic systems, and review different nonholographic and holographic approaches, taken by different researchers, in order to control the spatial and spectral characteristics of plasmonic waves. As will be thoroughly presented in the text, these holographic methods are described by the interference between two waves, a reference wave, which illuminates the hologram, and an object wave, which reconstructs the desired wavefront. As any of the two waves could either be a free-space wave or a surface-plasmon wave, four groups of holograms

Department of Physical Electronics, School of Electrical Engineering, Fleischman Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel
*Corresponding Author(s): e-mail: itayet@post.tau.ac.il, ady@post.tau.ac.il

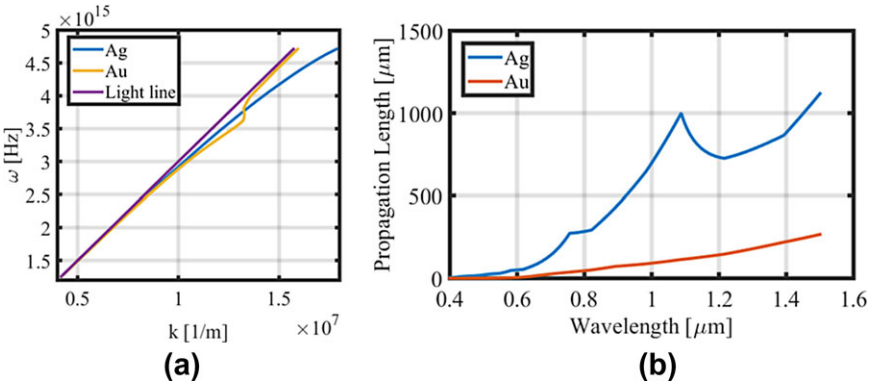


Figure 1 (a) Light line of air (purple) and dispersion relation of SPPs at a air/silver (blue) and air/gold (orange) interface. It can be seen that the wave-vector of the SPP, is larger than the wave-vector of the free-space wave, for any given frequency. (b) Characteristic propagation lengths for SPPs at air/silver and air/gold interfaces, in the visible and NIR range. (Ag and Au dielectric values taken from Johnson and Christy, "Optical Constants of the Noble Metals", Phys. Rev. B **6**, 4370–4379 (1972)).

could be identified – a free-space reference wave with a free-space object wave, a free-space reference wave with a plasmon object wave, a plasmon reference wave with a free-space object wave, and a plasmon reference wave with a plasmon object wave. Since our aim is to obtain a shaped plasmonic wave that is propagating in the near-field, and these are commonly excited with free-space light sources, we will focus in this review on the second type, e.g. the interference between a free-space reference wave with a plasmon object wave.

It should also be noted that while here we concentrate on shaping techniques of propagating surface-plasmon waves, there is also a lot of activity on shaping optical beams using localized surface-plasmons, or nanoantenna array-based metasurfaces [17–20], which will not be reviewed here.

The structure of this paper is as follows: we start by introducing the unique wave nature of surface-plasmons and basic concepts in holography. We then present the evolution of computer-generated holograms (CGH) and the challenges in achieving holographic control of plasmonic waves. This is followed by a comprehensive review of the different efforts, both nonholographic and holographic, conducted in order to control the plasmonic wave properties, which is followed by a comparison between these different methods through several examples together with their limitation. A summary is provided with the closing remarks.

2. Wave properties of surface-plasmons

The challenges in achieving full control over the wave properties of surface-plasmons reside in their unique and different wave properties, compared to those of free-space waves. These can be divided into the challenges relating to the optical excitation of SPP waves, and to the challenges in manipulating their wave properties.

The surface-plasmon-polariton (SPP) wave is a solution of the electromagnetic wave equation for a geometry of an interface between two semi-infinite layers, where one is a metal and the other is a dielectric. Assuming time-harmonic fields, the three-dimensional scalar Helmholtz equation is $\partial^2 E/\partial x^2 + \partial^2 E/\partial y^2 + \partial^2 E/\partial z^2 + k_0^2 \varepsilon_d E = 0$, where E is the electric field in the dielectric region, $k_0 = 2\pi/\lambda_0$ is the free-space wave-vector and ε_d is the dielectric permittivity, z is the propagation coordinate and x, y are the transverse coordinates, where y lies in the metal–dielectric surface and x is perpendicular to it. We define the metal/dielectric interface at the $x = 0$ plane. A wave of the form $E = A(y, z)\exp(ik_{sp}z)\exp(-k_x x)$ solves this Helmholtz equation, while maintaining the boundary conditions at the metal/dielectric interface. Only a transverse magnetic (TM) wave, for which the magnetic component of the electromagnetic field is nonzero in the Y–Z interface plane, satisfies the boundary conditions. $A(y, z)$ is the wave envelope, $k_{SPP} = k_0\sqrt{\varepsilon_m\varepsilon_d/(\varepsilon_m + \varepsilon_d)}$ is the SPP wave vector in the propagation direction, and ε_m is the permittivity of the metal. This wave exponentially decays in the x-axis in the dielectric region with a decay parameter $k_{x,d} = \sqrt{k_{sp}^2 - k_0^2\varepsilon_d}$. In addition, the wave is also decaying much faster in the metal region, with a decay parameter $k_{x,m} = \sqrt{k_{sp}^2 - k_0^2\varepsilon_m}$.

By performing the derivatives in x and using the expression for k_x, we can eliminate the x dependency from the wave equation and obtain the two-dimensional Helmholtz equation for the SPP envelope:

$$\partial^2 A/\partial y^2 + \partial^2 A/\partial z^2 + k_{sp}^2 A = 0, \tag{1}$$

with y being the only transverse coordinate.

For this SPP solution k_{SP} is complex, hence the wave also decays exponentially in the propagation coordinate$z$. This limits the propagation length of SPPs, which is defined as the distance at which the intensity drops to 1/e and is equal to $1/(2 \times \mathrm{imag}(k_{sp}))$. This inherent loss of the SPP wave is fundamental and acts as a major inhibiter for a variety of plasmonic applications. The absolute value of k_{sp} also determines the dispersion relation of SPPs. This solution therefore encapsulates the basic properties of the SPP solution – high spatial confinement to the metal/dielectric interface and loss of the optical field with propagation. Figure 1 shows the dispersion relation and characteristic decay lengths for SPPs at air/silver and air/gold interfaces. It can be seen that the propagation length is less than 100 μm in the visible, and is only a few hundred μm in the near-infrared.

Similarly to free-space waves, plasmonic solutions of the two-dimensional wave equation can be also studied

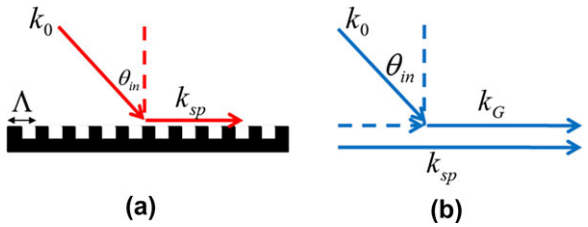


Figure 2 (a) Configuration for grating coupling of SPPs. A metal grating is illuminated at an angle and contributes momentum to the incident beam. (b) The corresponding momentum conservation diagram for the case of (a). In the horizontal direction, this diagram yields Eq. (3).

for the paraxial case, where the two-dimensional paraxial Helmholtz equation is:

$$\partial^2 A/\partial y^2 + 2ik_{\rm sp}\partial A/\partial z = 0. \tag{2}$$

This implies that both paraxial and nonparaxial plasmonic beams can be excited at the metal/dielectric boundary.

The dispersion relation of the SPP also yields that for any given frequency, the dispersion curve of SPPs lies to the right of the light-line for a given metal/dielectric interface [1], as seen in Fig. 1a. This means that the wave-vector of the SPP, $k_{\rm SP}$, is larger than that of the wave-vector of the free-space wave, k_0. Therefore, in order to excite an SPP wave from a free-space illuminating beam, an additional contribution of momentum to the free-space beam is required. There are several methods to achieve this additional momentum such as prism coupling, grating coupling, and others [1]. Here, we shall focus on one specific method that utilizes metallic periodic structures, namely diffraction gratings.

A one-dimensional (1D) square grating with period Λ, can contribute momentum via its grating wave-vectors, $k_{\rm G} = 2\pi m/\Lambda$, where the diffraction order m is an integer. Hence, when illuminated by a free-space beam the momentum conservation equation in the coupling direction yields (Fig. 2):

$$k_{\rm SP} = k_{\rm in} + k_{\rm G}, \tag{3}$$

with $k_{\rm in} = k_0 \sin\theta_{\rm in}$.

We note that this scalar definition of the grating neglects the effect of the ridge height on $k_{\rm G}$, which is often very small, considering that typical heights are much smaller than the wavelength.

In addition to the momentum compensation required to excite SPPs, one must also take into consideration that due to the limited propagation length of the SPP, the excited plasmonic wave should be measured directly in the near-field before it decays.

The above treatment applies for a monochromatic incident beam. For a spectral illumination of a broader nature, one should also take into account other factors. During SPP excitation by an incident free-space beam, spectral constituents for which both energy and momentum are conserved would efficiently couple to an SPP. Due to the grating's finite size, the momentum contribution is not single-valued, but rather a momentum distribution, centered about $k_{\rm G}$. Consequently, if the grating is not trivially a square grating in the propagation direction, it will affect the efficiency in which it couples certain wave numbers, when illuminated by a broad spectrum incident beam, enabling the possibility to spectrally shape the plasmonic excitation.

3. Holography of free-space waves

Holography is a general approach in wave phenomena, which enables recording and reconstruction of the full amplitude and phase information of a certain target, or object wave. This is achieved by recording the interference pattern of this object wave with a reference wave. Denoting the object and reference full waves (both amplitude and waves) as $A_{\rm o}$, $A_{\rm r}$, the intensity interference pattern is $H = |A_{\rm o} + A_{\rm r}|^2 = |A_{\rm o}|^2 + |A_{\rm r}|^2 + A_{\rm o}A_{\rm r}^* + A_{\rm o}^*A_{\rm r}$. Here, the first two terms of the right-hand side are proportional to the intensities of the object and reference wave, while the two last terms represent the interference of the two waves. These interference terms contain the desired amplitude and phase information of the object. This pattern can be recorded for example on a light-sensitive film. The wavefront can then be reconstructed by illuminating the hologram with the same reference wave. In this case we obtain $HA_{\rm r} = |A_{\rm o}|^2A_{\rm r} + |A_{\rm r}|^2A_{\rm r} + A_{\rm o}|A_{\rm r}|^2 + A_{\rm o}^*A_{\rm r}^2$. Here, the third term is the complex amplitude of the object. It has to be separated from the other three terms, and a method to achieve this is outlined below.

The original holography concept was proposed by Dennis Gabor in 1948 [21]. He was awarded the Nobel Prize in physics for this invention in 1971. Gabor's goal was to improve the resolution of electron microscopes. Electron holography is still widely used nowadays in electron microscopy to determine the phase variation of an electron beam, induced by electric and magnetic potentials [22]. In the early 1960s and following the invention of the laser, which provided an efficient method to realize the required reference wave with light waves, optical holograms became widespread.

The next important milestone was the invention of the computer-generated hologram by Brown and Lohman [23] in 1966. The key point is that instead of interfering a reference wave and an object wave on a light-sensitive film, it is possible to calculate directly the interference pattern and print it on a transparent slide. The reconstruction of the object wave is still performed in the same way of illuminating the hologram with a reference wave. This opens up many new possibilities, for example it enables creation of holograms of virtual objects that do not exist in reality, or to create dynamic holograms using computer-controlled spatial-light-modulators (SLM).

The holograms of Brown and Lohman [23] were binary amplitude holograms, i.e. the normalized transmission in each point of the hologram could take only two possible values either 0 or 1. To encode both amplitude and phase,

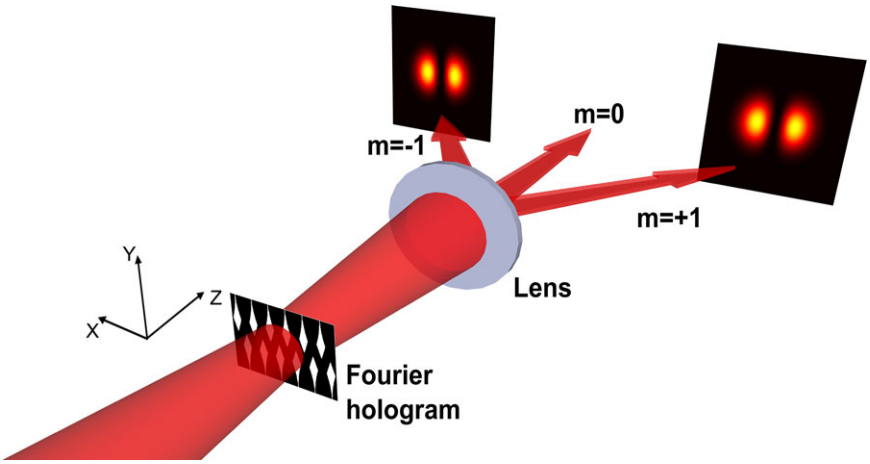


Figure 3 A Fourier hologram for far-field reconstruction. To obtain a desired beam shape in the far-field, in this case a Hermite–Gauss (1,0) beam, a thin hologram is encoded with the inverse Fourier transform of the target beam. The field emanating from the mask is then optically Fourier transformed by a lens, and generates the desired field in the first diffraction order. The different terms of the off-axis hologram are spatially separated in the far-field. Reproduced with permission [15].

they used the detour-phase method: The hologram is split into many small cells. In each cell there is a transmitting slit whose width determines the amplitude of the cell, and its location within the cell determines the phase.

While the hologram presented above directly encodes the wavefront, it is also possible to make a Fourier hologram that encodes the Fourier transform of the wavefront. In this case, the desired wavefront can be obtained by optical Fourier transform with a lens. If this is an off-axis hologram, in which there is a tilt angle between the object and reference wave, the desired wavefront can be easily separated from the three unwanted terms, namely those that are proportional to the intensities of the object and reference wave and the term that is proportional to the complex conjugate of the object wavefront. In the off-axis Fourier hologram, only the desired wavefront is obtained at the first diffraction order. Figure 3 shows the apparatus of such an off-axis Fourier hologram.

Three different schemes for binary modulation were proposed by Lee in 1979 [24]. These methods are based on converting the continuous interference term $A_o A_r^* + \text{c.c.}$ into a binary function. If one takes the sign of this term we obtain the phase information of the wavefront, but the amplitude information is lost. Lee suggested a clever way to digitize the interference term of a Fourier hologram, so that the amplitude modulation of the wavefront would also appear in the first diffraction order. As an example, one of Lee's methods takes advantage of the fact that the first order of the Fourier series of binary grating has an amplitude factor of $\sin(\pi q(x))/\pi$. This can now be used to determine the clipping level of the interference pattern. In order to make a Fourier hologram, we first compute the inverse Fourier-transform of the desired wavefront, denoted as $A_f(x)\cos[\varphi_f(x)]$, where the normalized amplitude is between 0 and 1 and the phase between 0 and 2π. A binary one-dimensional off-axis Fourier amplitude hologram is therefore defined as:

$$t(x) = 0.5\left\{1 + \text{sign}\left[\cos\left(\frac{2\pi}{\Lambda}x + \phi_f(x)\right) - \cos(\pi q(x))\right]\right\}, \tag{4}$$

where $q(x) = \sin^{-1}(A_f(x))/\pi$ and Λ is the period of the carrier frequency.

So far we have discussed holograms in which the reference waves and object waves are of the same type, e.g. both of them are optical free-space waves. In the following sections we will discuss plasmonic holograms, where there is a fundamental difference – since one of the waves is an optical wave, but the other one is a surface-plasmon-polariton wave [25]. As will be detailed below, this opens up interesting new possibilities, since it provides natural separation between the reference wave and object wave, and therefore enables reconstruction of the object directly in the near-field.

4. Fundamental challenges of surface-plasmon holography

As mentioned in Section 2, when trying to control the wave properties of SPPs, one encounters several fundamental challenges, owing to their unique wave nature. In order to fully control the plasmon wave properties with holographic methods, the following issues should be addressed:

First, and as previously discussed, coupling a SPP wave from a free-space wave requires a compensation for the missing momentum between the two wave-vectors. Secondly, in addition to the limited SPP propagation length there is also the issue of the limited measurement range of near-field characterization tools, such as near-field scanning optical microscopes (NSOM), emphasizing the need to both generate and detect the excited SPPs directly in the near field, before they decay. Thirdly, while free-space holograms are usually optically thin, thereby providing a well-defined plane of the amplitude and phase patterns for the illuminating free-space beam, SPPs are excited over a finite propagation distance and therefore their amplitude and phase cannot be simply defined at a specific one-dimensional plane. Fourthly, as mentioned in the previous section, most commonly used holograms for free-space waves generate additional wavefronts, or diffraction orders, to that of the desired beam, which can be separated from it in the far-field by a number of methods. For the case

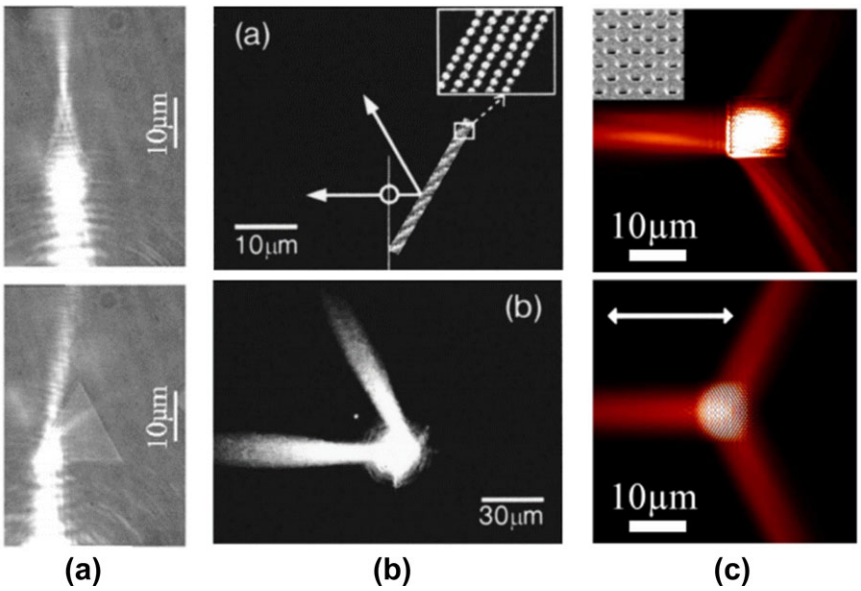


Figure 4 (a) Measurements of plasmonic focusing and reflection/refraction using dielectric features deposited on a metal surface. (b) Plasmonic mirror based on Bragg grating reflector. (c) Measured (top) and simulated (bottom) triangular plasmonic beam splitter based on hole-array manipulation. Reproduced with permission [26, 29, 30].

of SPPs this is a crucial problems since in the near-field all these wavefronts or diffraction orders will interfere together. Finally, at present there are no dynamic tools for controlling the wavefront of SPPs, like SLMs for the case of free-space beams.

Using holography for shaping the spectrum of SPP excitations can also come across difficulties, especially when trying to achieve broadband control together with a high excitation efficiency. A trade-off exists between the spectral efficiency and the hologram bandwidth, e.g. increasing the hologram's length in the propagation dimension will increase the coupling efficiency, however, it also reduces the bandwidth. This will be discussed in more detail in Section 8.

5. Nonholographic methods for spatial and spectral shaping of surface-plasmons

Early investigations of the spatial coupling and propagation characteristics of SPPs suggested that these can be shaped or controlled to a certain degree. These studies were based mainly on incorporation dielectric features on metal surfaces in order to create phase-retardation elements [26, 27]. In this method the equivalents of some basic optical elements were realized for plasmonic waves. For example, by depositing a cylindrical dielectric on a metal substrate, plasmonic lensing was demonstrated, and by depositing a triangular-shaped dielectric plasmonic reflection, refraction, and total internal reflection have been demonstrated [26] (Fig. 4a).

In a different approach, plasmonic planar Bragg gratings were realized by the deposition of dielectric spherical scatterers on a substrate, followed by a metal deposition [27], or based on dielectric gratings on a metal surface [28]. In this method, Bragg reflectors were realized to fold the plasmonic beam similar to optical mirrors, together with beam splitters and interferometers (Fig. 4b). Another approach utilized the passive and even active manipulation of two-dimensional metal hole arrays [29–31]. For example, by spatially shaping the array, several plasmonic schemes for focusing have been realized, together with triangular beam splitting [29] (Fig. 4c).

A significant driver to the generation of more complex plasmonic beams was the introduction and demonstration of the optical Airy beam in free space [32, 33]. The Airy beam is a unique solution of the paraxial Helmholtz equation (Eq. (2)), which exhibits unusual properties; it is a "nondiffracting" beam – i.e. it preserve its spatial shape with propagation. It is also a "self-healing" beam – i.e. it can reconstruct itself to its original shape after being distorted by obstacles, and above all, it is a "self-accelerating" beam – i.e. it propagates along curved trajectories. It was suggested in 2010 that the unique properties of the Airy beam would also be supported in a two-dimensional system as a surface wave, e.g. the Airy plasmon [34, 35]. This led to the first generation of a plasmonic Airy beam in 2011 by three different groups, and in three different methods. Minovich et al. [8] relied on the realization of a kind of Airy grating coupler, e.g. a special grating that closely imitates the Airy-beam characteristics, and then improved using optimization methods in order to obtain the an Airy plasmon beam. The relation between the Airy grating coupler and the phase and amplitude of the Airy function, together with the experimental apparatus, can be seen in Fig. 5a. Li et al. [9] relied on the inplane diffraction method, which is based on a coupled SPP that diffracts from an array of scatterers. By controlling the spatial arrangement of these scatterers, the phase front of the scattered SPP can be designed for different applications [9, 12, 36, 37] (Fig. 5b). Using these methods for the generation of a SPP with a spatial phase front possessing a 3/2-power phase modulation lead to the generation of an Airy-like plasmon [9]. Zhang et al. [10], realized the Airy plasmon by generating a free-space Airy beam, which was then coupled to an Airy plasmon using a regular 1D metal grating (Fig. 5c).

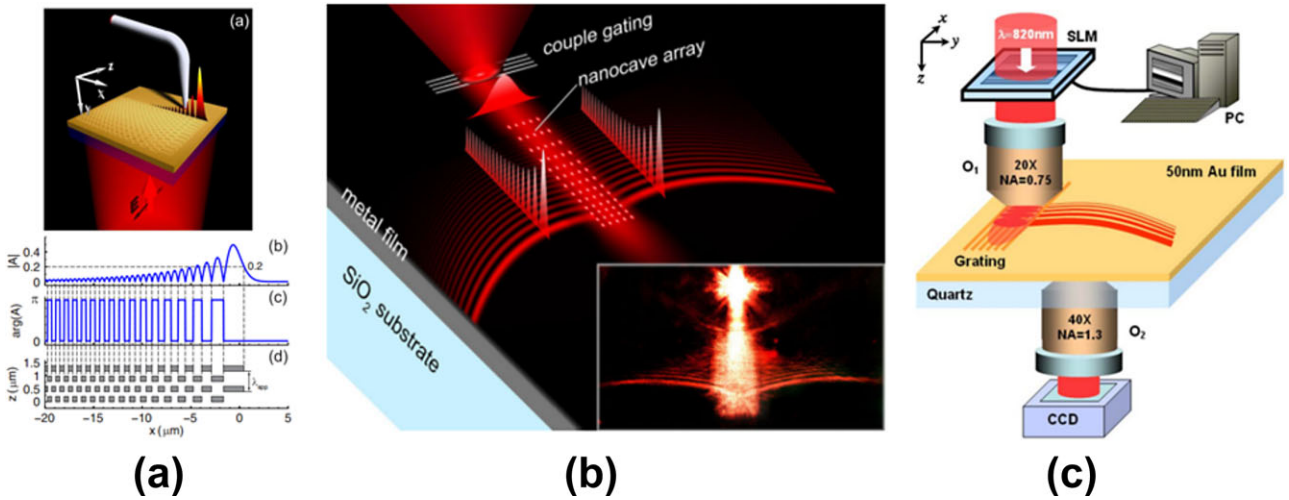


Figure 5 Three methods for generating plasmonic Airy beams. (a) An optimized grating coupler, build out of rows of square holes milled in a gold layer, produces an Airy beam output. (b) An array of scatterers, obtained by the inplane diffraction method, forms a plasmonic Airy beam. (c) A free-space Airy beam is coupled to an Airy plasmon beam using a regular periodic grating. Reproduced with permission [8–10].

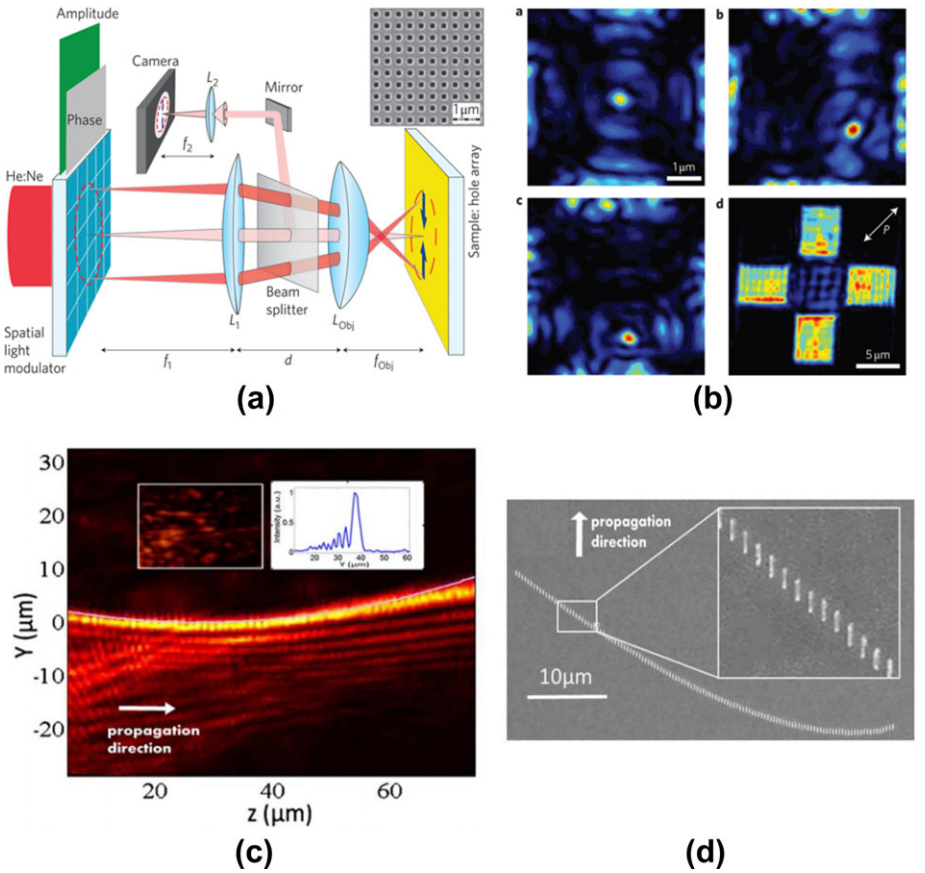


Figure 6 (a) Optical setup for dynamic control of SPPs. An SLM is imaged on a subwavelength hole array and the image is collected by a camera. (b) Generation of standing waves and focal spots that can be dynamically controlled. (c) Surface-plasmon caustic generated by the nanoantennas array presented in the SEM image in (d). Reproduced with permission [38, 57].

Another approach for dynamically controlling the spatial properties of plasmonic waves is by the use of an SLM [38, 39]. In this method, the SLM is imaged onto a metallic hole array, illuminating it partially, in order to excite SPP only at those illuminated areas (Fig. 6). Exciting oppositely propagating SPPs results in an interference pattern and in this way standing waves and focal points can be realized and even be used for microcopy [39]. The use of the SLM enables dynamical control of these patterns and the location of the focal points.

Along with the efforts to spatially control the excited plasmonic wavefront, applications of SPPs for

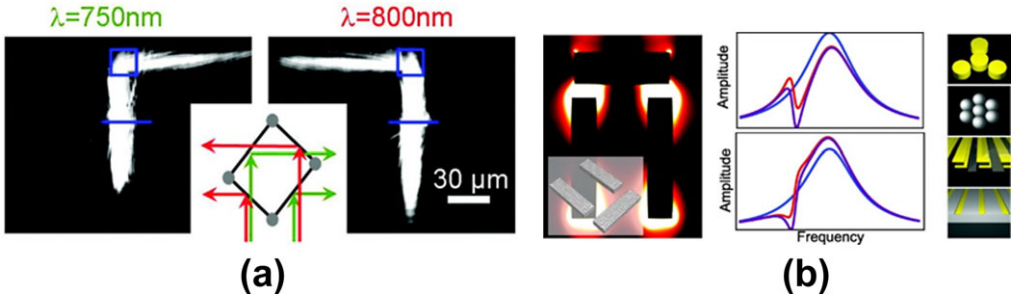


Figure 7 Two applications for plasmonic spectral shaping: (Left) Implementation of the plasmonic demultiplexer, which separates an incident SPP beam into two 180° separated spectral channels. (Right) Shaping the spectral Fano-like response of metastructures having applications in refractive-index sensing. Reproduced with permission [40, 58].

communication and for onchip circuitry also necessitate the control over the excitation energy spectrum. One of the first, and perhaps the most desirable spectral device is the demultiplexer [40], seen in Fig. 7a, which allows for sorting polychromatic light beams into separate SPP waveguides, or beams. Succeeding efforts were made to launch unidirectional, broadband SPPs using various methods such as magnetic antennas [41] and deep-subwavelength slits in substrate illumination [42]. These solutions had succeeded in coupling a broadband SPP wave, however, they did not provide sufficient degrees of freedom for completely controlling the excitation spectrum. Experiments exhibiting control of the plasmonic spectrum by shaping an ultrashort pulse incident on a patterned tip have also been recently reported [43]. This route seems to be promising as it might enable dynamically changing the SPP excitation spectrum.

As mentioned earlier in the introduction, there has been in recent years a growing interest in shaping optical beams using localized surface-plasmons, or nanoantennas, which are referred to as metasurfaces [17–20]. These are aimed to shape the optical beam at the far-field, rather than the near-field, and owing to this are beyond the scope of these paper. However, a few demonstrations have also been made in order to use these nanoantenas to shape propagating surface plasmons in the near-field, due to their ability to couple light to surface plasmons in a localized manner. Such localized excitation techniques have been suggested using a near-field probe [44] and electron-beam excitation [45]. Another approach is to use subwavelength metallic nanostructures that serve as scattering centers and couple light to SPPs [46, 47]. The nanostructures have a spectra response that can be tailored for a desired wavelength and polarization by designing their geometry. At the relevant wavelength a resonant response is obtained and the scattering is of high efficiency. In this term they can efficiently convert far-field radiation to near-field currents, and vice versa, and are known as optical nanoantennas [48]. A variety of schemes for coupling based on such metallic nanostructures are available [49–53], including functional beam shaping [54]. These methods demonstrate the advantages of using nanostructures to manipulate light and SPPs with relatively high efficiency, which ranges from about 1–3% [55], compared to a previously estimated 0.35% for gratings [56], and in a compact geometry. By spatially arranging these nanoantenas on a metallic surface, the unique properties of localized surface-plasmon excitation have been utilized to shape the excited surface-plasmon beam. The nanoantennas were designed to excite propagating surface-plasmons on the metallic surface and their spatial arrangement on the surface is designed to launch the propagating surface-plasmon with a specified wavefront. Since the spacing between the nanoantennas can be subwavelength, the resulting plasmonic wavefront is relatively smooth and uniform. In this manner, surface-plasmon curved caustics, bottle beams and dual-focii lenses were demonstrated [57]. Figure 6 presents an example of a surface-plasmon caustic (Fig. 6c) that was generated by an array of silver nanoantennas spatially arranged on a silver surface in a manner that manifests the phase required to generate the plasmon caustic wavefront (Fig. 6d).

We note that significant progress has also been made in tailoring the spectral response of localized SPPs [58], seen for example in Fig. 7b.

The examples above provided significant steps towards controlling the plasmonic wavefront and its spectrum, and also implied that a more general approach in order to fully control the plasmonic wave properties would be highly beneficial. One such approach, which is discussed in the next section, is the use of holographic methods.

6. Holographic methods for spatial and spectral shaping of surface-plasmons

The concept of holography has also been extended to plasmonic waves. For example, interference between a plasmonic reference and a plasmonic object beam was suggested, in order to define lines that can then be etched in the metal surface. This forms a plasmonic hologram that can convert the reference wave into a desired reconstructed wave [59, 60] (Fig. 8a). This elegant method assumes that the plasmonic reference wave is already present, and does not address the coupling issue of the interacting plasmons. This approach, which is in principle independent of the coupling process, would be of great advantage for compact photonic integration.

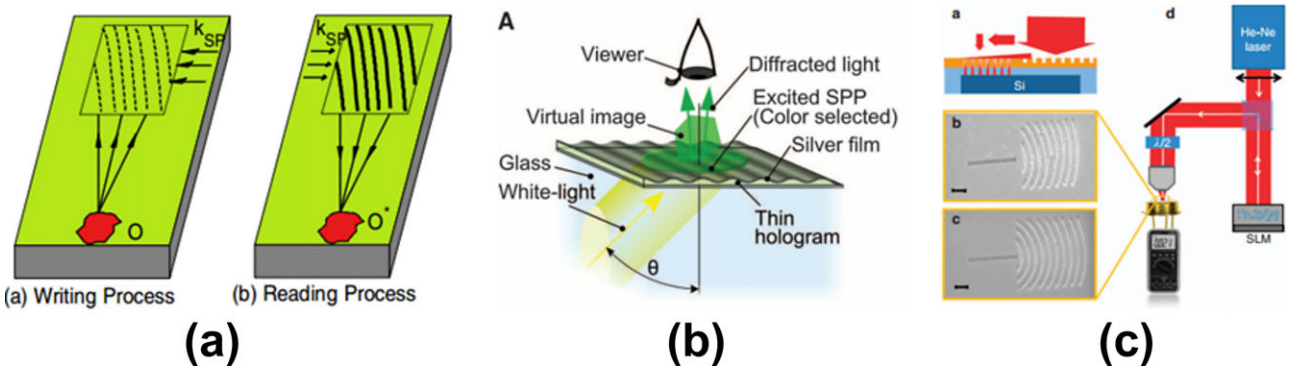


Figure 8 Holographic methods implemented with SPPs. (a) a read-and-write process with both the object and reference beams are plasmonic beams. (b) A read-and-write process with a reference plasmonic beam and a free-space object beam. (c) Similar approach as in (b) but by using CGH rather than a read/write process. Reproduced with permission [60, 62, 67].

In another scheme, interference between a reference plasmonic beam and an object free-space beam was used to reconstruct a desired free-space beam wavefront at the far-field. The latter was excited by a plasmonic beam which diffracted from the hologram [25, 61–65] (Fig. 8b). A similar approach, where a plasmonic reference beam was used to generate an object free-space beam, was applied using CGH. This approach eliminates the need for interfering both beams in order to obtain the hologram's pattern. This involves the use of two plasmonic gratings, the first couples a free-space beam into a plasmonic beam, which then propagates to the plasmonic CGH, which in turn couples the plasmonic beam back to free-space with the required amplitude and phase. This method was previously used to couple plasmonic beams with free-space Airy and Laguerre–Gauss vortex beams [66]. In the reversed approach, free-space vortex beams, carrying a certain angular orbital momentum, corresponding to a certain topological charge, were shined on a plasmonic hologram also possessing a different topological charge [67] (Fig. 8c). This resulted in a change of the total topological charge of the generated plasmonic beam, which affects its spatial shape. From the latter, one could detect the original orbital angular momentum carried by the free-space beam.

These methods were designed to control or detect the free-space wavefront, generated in the far-field, using the surface-plasmon beam attributes that were located in the near-field. In the next section we will discuss the introduction of another class of plasmonic holograms, which are specifically designed for controlling the near-field wavefront of surface-plasmon waves.

6.1. The binary near-field hologram

In order to obtain holographic control over the amplitude and phase of the plasmonic wavefront directly in the near-field, we can similarly define the plasmonic hologram as an interference between two beams. However, this time the reference wave will be defined as the free-space beam and the object wave will be defined as the plasmonic beam. Following this approach and similarly to the Fourier hologram described in Section 3, the desired plasmonic beam will be generated in the first far-field diffraction order of the plasmonic beam, e.g. one needs to propagate the plasmonic beam to its "far-field" in order to obtain the desired beam's shape. Unfortunately, such propagation to the Fraunhofer region without using Fourier-transforming plasmonic lenses (Fig. 4a) is impractical due to the limited propagation lengths of SPPs. Moreover, several additional plasmonic beams will be generated at all the other diffraction orders of the hologram, making this approach inefficient.

A robust approach to overcome this issue resides in a different class of holograms, which are designed specifically for the near-field [15]. In this near-field hologram (NFH), the free-space beam acts as the reference beam, while the plasmonic object beam is obtained directly in the near-field, rather than the far-field. This is feasible due to the fact that there is a Fourier transform (FT) relation between the near-field and the far-field. In Section 3 we discussed the principle of Lee's binary encoding method, in which the inverse FT of an object beam was encoded into the hologram in order to obtain it in the first diffraction order in the far-field. This means that there is also a FT relation between the field emanating from the hologram, e.g. the near-field, and the field obtained in the first diffraction order in the far-field. Thus, if one should desire to obtain the object field in the near-field, one can actually encode into the hologram the object field itself, rather than its inverse FT, as was done in Section 3. The latter is what makes the NFH so useful for plasmonic waves – it enables us to obtain the object plasmon wave directly in the near-field, right at the vicinity of the hologram, where it can be detected before decaying.

In the scheme of the plasmonic NFH, the NFH itself lies in the plane of propagation, unlike that of the Fourier hologram, and is perpendicular to the direction of the free-space beam illumination, as shown in Fig. 9. The plasmonic object field, namely its amplitude and phase $A(z, y)$, $\phi(z, y)$, are encoded along the transverse coordinate $y$with a modulated spatial carrier frequency of $2\pi/\Lambda$ in the $z$coordinate, which is the propagation direction, where Λ is the period of

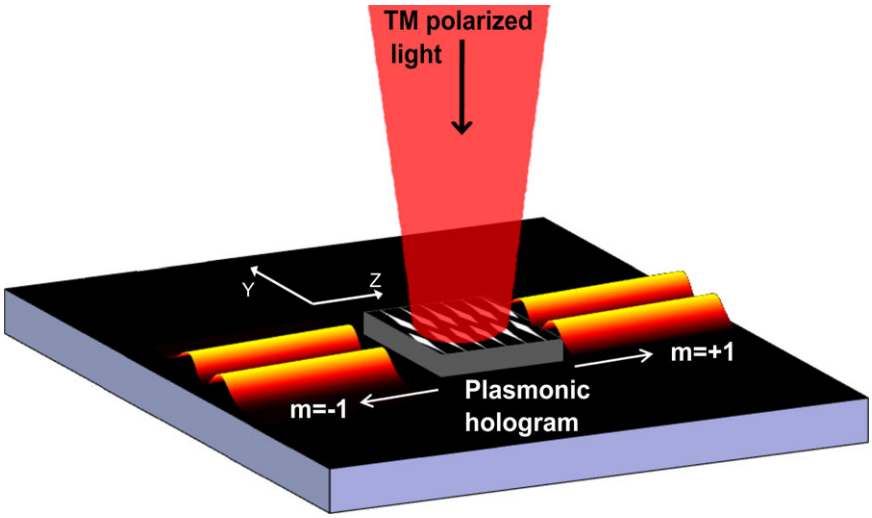


Figure 9 The scheme for a plasmonic near-field hologram. The hologram lying in the (z, y)-plane, is encoded with the direct amplitude and phase of the desired beam. The hologram is then illuminated by a free-space beam, and its first diffraction order couples into a SPP. The generated plasmonic field by the hologram in the near-field is the desired field, while the other terms of the hologram are being diffracted back to free space. Reproduced with permission [15].

modulation. This modulation acts as a grating in the propagation direction. Therefore, by choosing the appropriate period Λ, one can use this grating to contribute the missing momentum between the SPP and free-space wave-vectors, as discussed in Section 2, thus coupling it to the interface. In this way, one inherently gains the separation of the desired first diffraction order of the hologram from all the others, namely, only it will couple to an SPP since only it conserve the required momentum. Specifically, for zero-angle illumination, as illustrated in Fig. 9, the $+1$ and -1 orders of the hologram will propagate in opposite directions from the hologram along the interface. Moreover, for a nonzero angle, all diffraction orders except the coupled $+1$ order will diffract back to free-space. This is a major advantage of the plasmonic NFH compared to the Fourier hologram, since as mentioned before, when using a Fourier hologram the different diffraction orders cannot be separated in the near-field.

Owing to the binary nature of metallic gratings used to couple SPPs to the metal/dielectric interface [1], we can use a modification of Lee's binary encoding method that was introduced in Section 3, in order to obtain the two-dimensional plasmonic NFH using the following equation [15]:

$$t(z, y) = \frac{h_0}{2}\left\{1 + \text{sign}\left[\cos\left(\frac{2\pi}{\Lambda}z + \phi(y)\right) - \cos(\pi q(y))\right]\right\}, \tag{5}$$

where $q = \sin^{-1}[A(y)]/\pi$ and contains the amplitude information, Λ is the period of modulation and h_0 is the hologram's height. Equation (5) encapsulates all the required attributes of the plasmonic hologram: an arbitrary plasmonic field, $A(z, y)$, $\phi(z, y)$, can be encoded, a carrier frequency completes the missing momentum, thus coupling only the first-order SPP from free-space beam to the interface, and the desired beam is generated directly in the near-field. One can also see from Eq. (5) that the modulation is encoded in the direction of propagation z, while the amplitude and phase of the desired beam is encoded in the transverse direction y. Another difference from Lee's original code (Eq. (4)) is that rather than using the Fourier transform of the desired wavefront to form a Fourier hologram, here the wavefront is encoded directly, as discussed above. This approach can now be readily used to generate any desired plasmonic beam for a variety of applications.

6.2. Case study – high-order Hermite–Gauss plasmon modes

In order to provide more intuition to the reader on how and why this plasmonic NFH works, we present here a step-by-step example, for the use of a plasmonic NFH in order to generate a Hermite–Gauss (HG) plasmonic beam, which is one of the most basic beams used in optics. HG modes are solutions of the paraxial Helmholtz equation (Eq. (2)), and exhibit the self-similarity property, meaning that they diffract with propagation but maintain the same scaled beam profile. Owing to the dimensionality of plasmons, the plasmonic HG beam is one-dimensional, and described by the general one-dimensional, high-order HG mode formula:

$$u_{m,0}(y) = \frac{A_m}{w_0} H_m\left(\frac{\sqrt{2}y}{w_0}\right) \exp\left(-\frac{y^2}{w_0^2}\right) \tag{6}$$

where w_0 is the beam radius and H_m are the Hermite polynomials.

For our case study we choose to generate the HG(1,0) mode, and as a first step we isolate its amplitude and phase from Eq. (6). Figure 10 shows the normalized absolute amplitude and phase of the HG(1,0) mode. We can now insert the obtained amplitude and phase to Eq. (5) in order to obtained the spatial distribution of the plasmonic NFH for the HG(1,0) mode. A SEM image of this NFH fabricated from silver metal is shown in Fig. 10b. Examining the relation between the object HG(1,0) field presented in

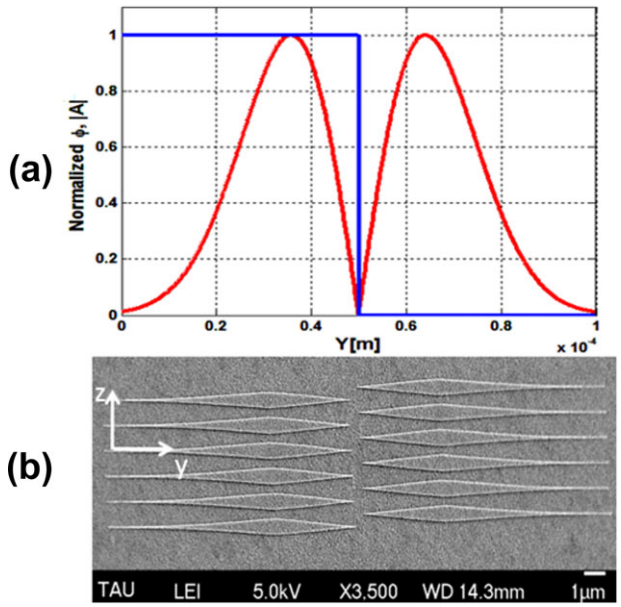


Figure 10 (a) Normalized absolute amplitude and phase of the HG(1,0) mode, and SEM image of the corresponding plasmonic NFH (b).

Fig. 10a and its resulting NFH in Fig. 10b, reveals the nature in which the information is encoded into the NFH. It can be seen that in the propagating direction$z$the pattern repeats itself with periodΛ, corresponding to our choice of modulation period, and this creates a grating structure. We can observe that the peaks presented in the amplitude of the HG(1,0) mode correspond to a duty cycle (DC) of 50% in the NFH (black arrow), which correspond to the highest coupling efficiency of the NFH. Furthermore, we can see that the decrease in amplitude towards the edges of the HG(1,0) mode correspond to a decrease in the DC towards the edges of the NFH (black arrow). It is now clear that the DC of the NFH is a manifestation of how the amplitude of the object beam is encoded into the NFH. We can also observe that the left part of the NFH is shifted relative to the right part by exactly half of the modulation period. This shift exactly corresponds to the π phase shift seen in the relative phases between the left and right lobes HG(1,0) mode. This spatial shift in the NFH is a manifestation of how the phase of the object beam is encoded into the NFH.

The last step remaining is to see what will be the plasmonic field that will be generated by this NFH. Measurements of the plasmonic field intensity for two different HG mode, HG(1,0) and HG(3,0) plasmonic beams, are shown in Fig. 11, together with SEM images of the plasmonic NFHs that were used to generate them. The inset shows a comparison between the theoretical and measured beam profile for both beams, and these differences will be discussed in detail later.

Although the HG example presents both amplitude and phase encoding of the object field, it is sometimes desirable to realize phase-only elements. In this case, there is no need to encode the amplitude and the $\cos(\pi q(y))$ term in Eq. (5) can be omitted. As an example of a plasmonic phase-only NFH we present the generation of a plasmon caustic. A caustic is a curved light beam that is constructed from many geometrical rays that are tangent to the trajectory of curvature. There is a specific relation between the curvature of the beam and its initial phase, therefore, in order to generate a caustic one need to realize its relevant phase profile. Both paraxial and nonparaxial arbitrary trajectory caustics have been realized in free-space using this method [68, 69]. This means that in order to generate a plasmon caustic one needs to realize this relevant phase in the near-field, and this can be done with a phase-only

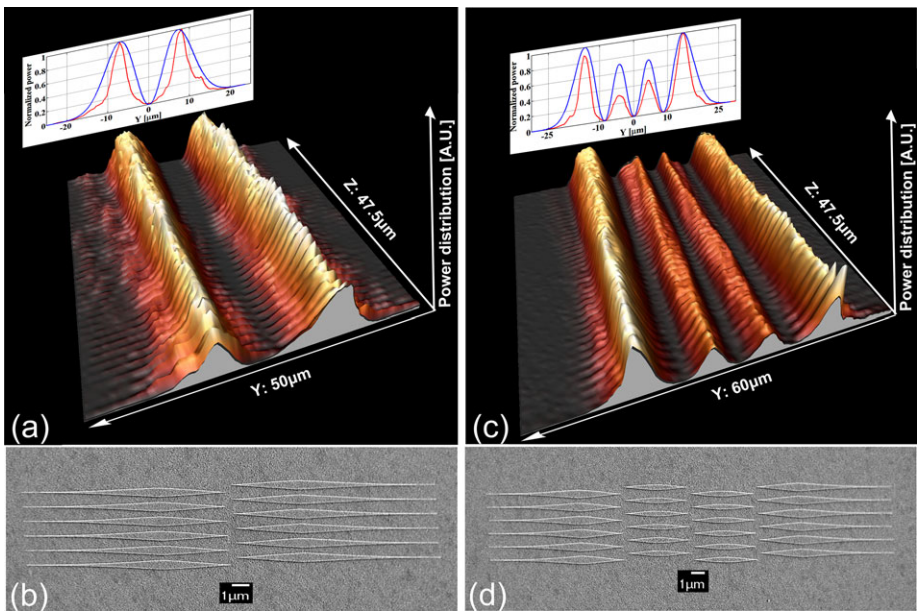


Figure 11 NSOM experimental measurement of plasmonic (a) HG(1,0) and (c) HG(3,0) beams. The insets show the cross section of the simulated (blue curve) and measured (red curve) intensities. SEM images of the plasmonic holograms which generated the HG(1,0) and HG(3,0) beams, are shown in (b) and (d), respectively. Reproduced with permission [15].

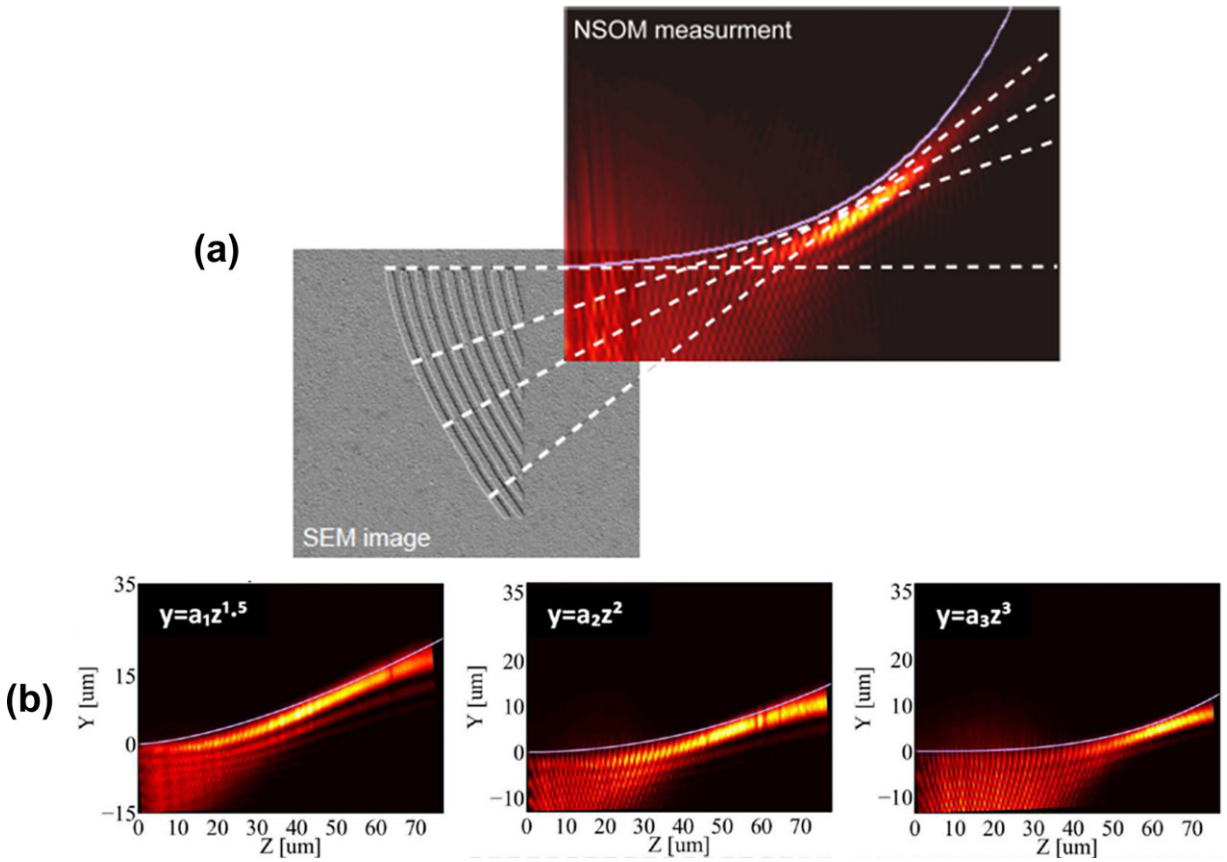


Figure 12 (a) Geometrical representation of the construction of a caustic SPP. Geometrical rays (white dashed lines) emanating from the two-dimensional plasmonic binary phase mask (SEM image) which generated the caustic SPP (NSOM measurement). The mask and the measured SPP are for the case of an exponential trajectory (solid purple line shows the analytical curve). (b) Experimental demonstrations of plasmonic SAB with three different polynomial trajectories. Reproduced with permission [13].

NFH. Figure 12 shows the realization of several plasmon caustics by a plasmonic phase-only NFH. Figure 12a offers a graphical representation of the construction of the caustic SPP, for the case of an exponential trajectory, and it can be seen that the plasmonic NFH generates geometrical rays at different angles that are tangential to the constructed exponential caustic SPP curve. It is therefore understood that the curvature along the direction $y$of the NFH is a manifestation of the required caustic phase. Similarly, additional arbitrary caustic trajectories can be realized by implementing different phase-only NFHs [13] and Fig. 12b presents three examples of different polynomial trajectories.

6.3. Simulating the propagation of plasmonic waves using Green's function

It would be highly beneficial to simulate the target object beam and the holographically generated plasmonic waves before the hologram fabrication is performed. Due to the nature of the hologram being a 3D structure, it is customary to use three-dimensional FDTD or FEM software in order to simulate the plasmonic field that will be generated by the mask. These are usually quite expensive, complex to use, and require a long running time owing to the 3D nature of the simulation. For a comprehensive review of numerical methods to simulate nanophotonic devices, see [70]. We note that for beam-propagation applications, as required in the case of plasmonic beam shaping, we are only interested in the 2D scalar propagation attributes of the plasmonic field, as we know that the behavior perpendicular to the interface will exhibit an exponential decay into both metal/dielectric materials. Therefore, it will be beneficial to find a simpler two-dimensional approach.

Since SPPs are TM solutions of Helmholtz's equation, their propagation dynamics at the plasmonic metal/dielectric interface can be described by the 2D scalar Helmholtz equation, Eq. (2). Here, we assume that all SPPs that have been successfully excited at the interface are TM, and from here the scalar treatment can hold. Therefore, we can now treat the propagation dynamics as being governed by the 2D Green's function of the 2D Helmholtz equation, which is $G_{2D} = -\frac{i}{4}H_0^{(2)}(kr)$, where $H_0^{(2)}(kr)$ is the Hankel function. The physical meaning of this treatment is that a 1D wavefront can now be treated as a collection of point sources, emitting 2D circular waves, similarly to the Huygens–Fresnel principal in 3D. The field at any given point $g(y, z)$ is a convolution of the sources

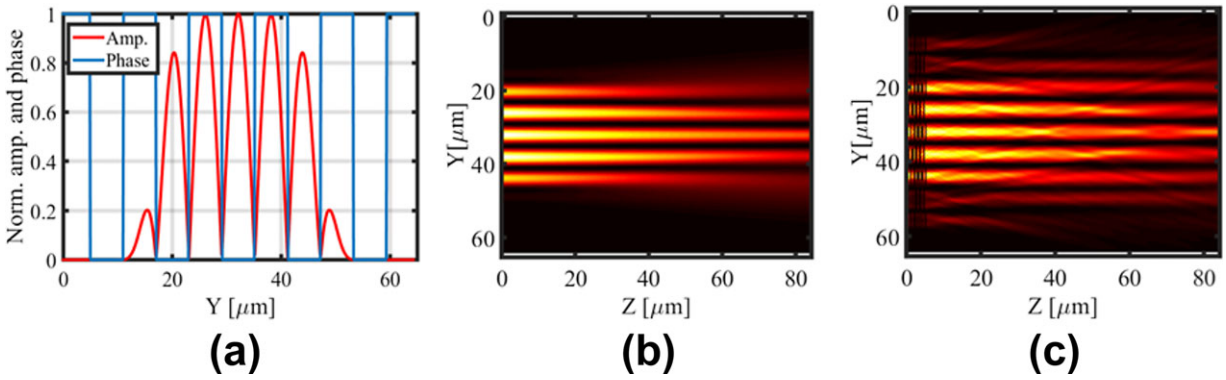


Figure 13 (a) Theoretical normalized amplitude and phase of a cosine-Gauss beam. (b) The theoretical cosine-Gauss beam obtained by propagating the field in (a). (c) The generated cosine-Gauss beam, obtained by propagating the field emanating from the plasmonic hologram, shown in the first 5 μm of the figure. The difference between (b) and (c) can be clearly seen and are due to the diffraction from the finite features of the hologram.

distribution $f(y, z)$ with the appropriate Green's function, $G(y - y', z - z')$:

$$g(y, z) = \iint f(y', z')G(y - y', z - z')dy'dz'. \quad (7)$$

To simulate the propagation of a desired arbitrary plasmonic beam, one has only to insert into this calculation the known 1D phase and amplitude profile, at the starting line z = 0, and perform the overall contribution of all these point sources over the 2D propagation space (Fig. 13). Moreover, since the treatment is that of a collection of point sources, this approach is not limited to a 1D profile, but also holds for 2D point-source distributions. This fact enables us to predict the output of the plasmonic holograms in a very precise way. The hologram itself can be divided into discrete sources, and therefore treated as a collection of 2D points sources. In a similar way, the contribution from all these sources can now be calculated over the entire 2D propagation space. This calculation can easily be implemented and Fig. 13 shows how this simple approach can be used to predict the output of plasmonic holograms very accurately. In addition, this method can also be used to simulate a non-normal incident illumination beam by adding a phase to the sources.

6.4. Fabrication of binary near-field holograms

The fabrication of the plasmonic near-field holograms is based on standard nanofabrication processes. A substrate is first coated with a thick metal layer, usually silver, of around 200 nm thickness. It is then spin coated with an electron-beam resist layer, such as PMMA [polymethyl methacrylate], which is followed by standard electron-beam lithography of the hologram's pattern onto the PMMA. After the chemical development of the PMMA resist, another metal layer is deposited on top of the PMMA pattern, with a thickness depending on the metal and wavelength of operation, for silver and 1064 nm, this thickness is 40 nm. Then, a lift-off process is performed to remove the undesired metal regions and the PMMA (see Fig. 14 for the fabrication steps). The plasmonic near-field holograms, presented in the next chapter, were designed for free-space illumination at normal incidence, at which the surface-plasmon wavelength is equal to the period of the spatial carrier frequency. In general, these can be designed to be illuminated by any desired angle, by changing the modulation period to satisfy the wave-vector equation, as illustrated in Fig. 2.

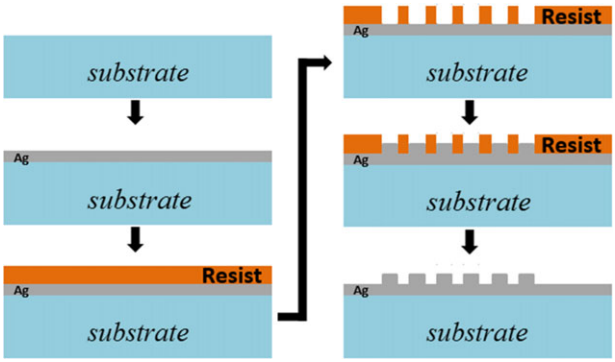


Figure 14 Illustration of the fabrication steps required for the realization of near-field plasmonic holograms.

7. Comparison between the different methods of SPP spatial shaping

So far we have discussed different approaches and methods for controlling the plasmon wavefront, being either holographic or nonholographic. It would be beneficial to perform a comparison between these different approaches in order to gain more insight into the advantages and disadvantages of each method. For that purpose, in this section we will review the experimental results, obtained by different methods, for the generation of several common

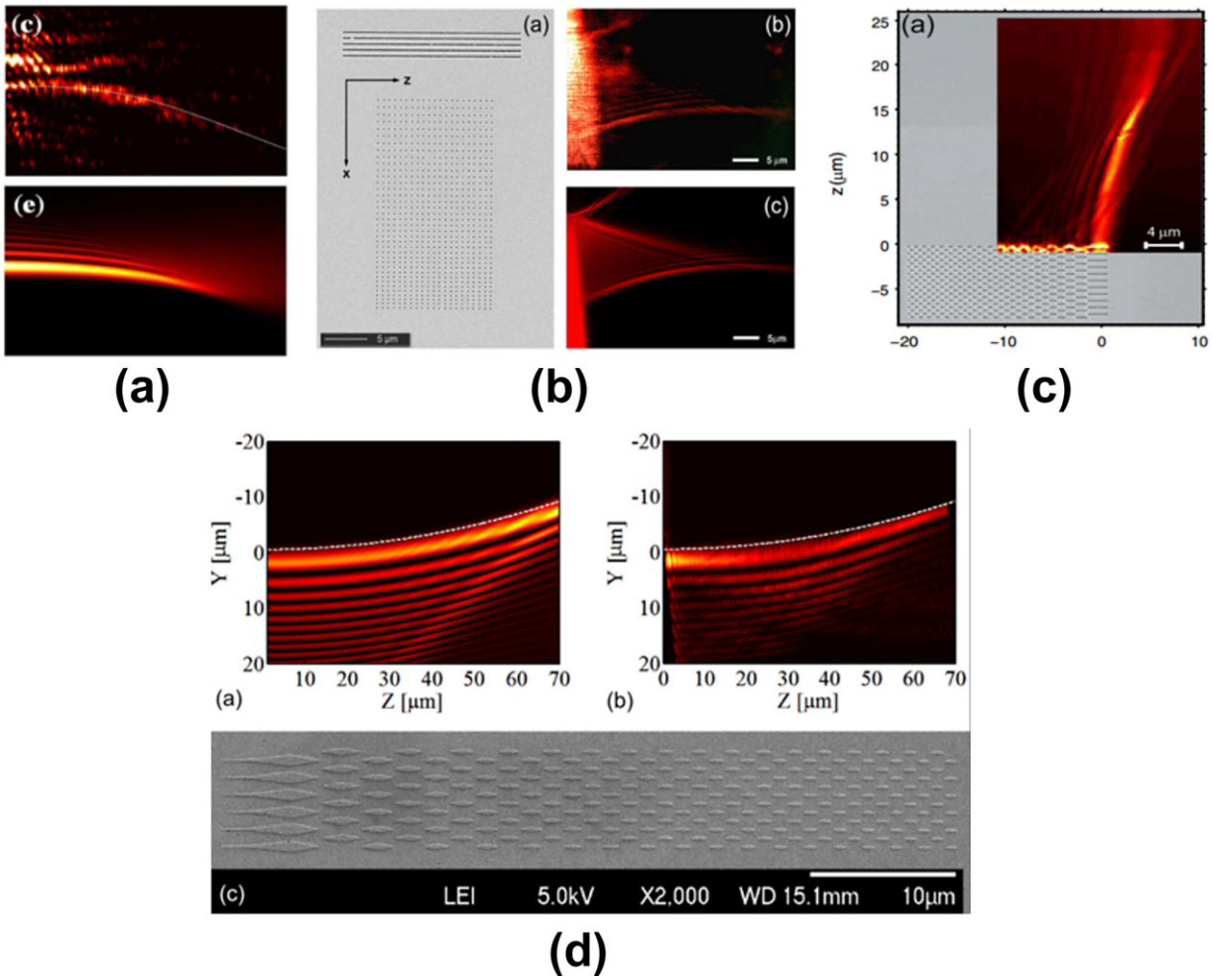


Figure 15 (a) Measured (top) and simulated (bottom) Airy plasmon using the method of Zhang et al. [10]. (b) Measured (top right) and simulated (bottom right) Airy plasmon using the method of Li et al. [9] together with the scattering array (left). (c) Measured Airy plasmon using the method of Minovich et al. [8] together with the Airy grating coupler. (d) Measured (top right) and simulated (top left) Airy plasmon using the plasmonic NFH method of Epstein and Arie [15] together with s SEM image for the NFH itself (bottom). Reproduced with permission [8–10, 15].

plasmonic beams. Specifically, we will discuss the cases of the plasmonic Airy, cosine-Gauss and Bessel beams.

7.1. Plasmonic Airy beam

We start by examining the different methods for generating a plasmonic Airy beam. As mentioned in Section 5, three nonholographic methods have been previously used to generate the Airy plasmon, together with a holographic generation of the Airy plasmon, based on the plasmonic NFH [15]. Figure 15 presents the results, and the structures used in each method, for the Airy plasmon generation.

The free-space Airy beam realization of Zhang et al. [10] (Fig. 15a) followed by its coupling to an SPP Airy plasmon by a grating, allows for dynamic control of the trajectory of the generated Airy plasmon. This is achieved by dynamically changing the phase of the free-space beam using a SLM, which in turn affects the trajectory of the coupled Airy plasmon. This approach, however, can work well only for paraxial beams, such as the Airy beam, as this kind of linear coupling relation, between the free-space beam and the plasmonic beam, can only hold under the paraxial approximation [71]. For generation of more complex and nonparaxial beams, this coupling approach requires a more careful structuring of the properties of the illuminating free-space beam in order to obtain greater control on the generated plasmon [71].

The inplane diffraction method, introduced by Li et al. [9] allows for a greater amount of control on the phase of the generated SPP, the ability to handle a broader spectral SPPs excitation [35], and is not limited by the above-mentioned coupling condition. This fact makes it more versatile for a variety of applications. However, owing to its use of scatterers, it is less efficient, especially when considering that the coupled SPP first propagates along the scattering array, and forms the generated SPP only when scattered from the array itself, symmetrically on both sides of the array.

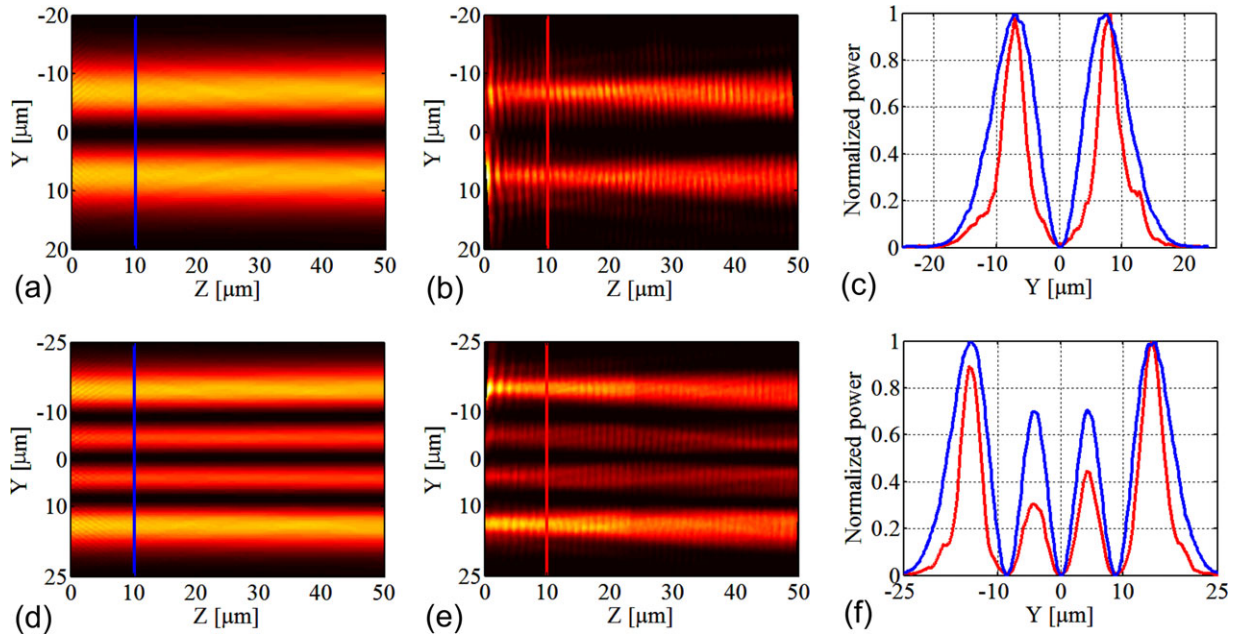


Figure 16 (a) and (d) numerical simulations, (b) and (e) NSOM measurements, and (c) and (f) their cross section at $z = 10\ \mu$m (simulated – blue curve, measured – red curve), of HG(1,0) and HG(3,0) plasmonic beams, respectively. Reproduced with permission [15].

Due to the propagation losses of SPPs this also decreases the propagation distance of the generated plasmonic beam. The generation of the Airy plasmon based on the inplane diffraction can be seen in Fig. 15b. Since the inplane diffraction method is a phase-only method, one should note that the generated beam is Airy-like and not the analytical Airy plasmon solution, as it is based on the generation of a 3/2-power phase modulation in the Fourier plane and not the analytical Airy solution composed of both phase and amplitude modulation.

The latter issue has been addressed by the Airy grating coupler demonstrated by Minovitch et al. [8], shown in Fig. 15c. Examining a single line of the Airy grating reveals its incorporation of the Airy amplitude modulation. As the Airy beam's amplitude is composed of gradually decreasing spatial lobes (Fig. 5a), the line is composed of gradually decreasing spatial rectangular slits, which only couple free-space light to SPP where the slit exists. Since the Airy beam's phase is composed of alternating π phase shift between adjacent lobes the second line of the grating generates the complementary amplitude modulation of the first line, but spatially shifted by half the wavelength of the SPP. These two lines then repeat themselves periodically to form the complete Airy grating. The resulting Airy plasmon can be seen in Fig. 15c as well, and although this approach incorporates amplitude and phase modulations, it still does not provide the generation of the analytical Airy beam solution. This is due to the fact that the exact Airy beam lobes' intensity gradually decrease in intensity as well and have a smooth round shape. These are not addressed by the use of square slits in different sizes, and additional optimizations are required in order to generate the resulting beam.

In [15], an Airy plasmon solution with the analytical amplitude and phase dependence was generated by a realization of a binary plasmonic NFH, according to the analytic Airy function. Measurements of the generated Airy plasmon and a SEM image of the Airy plasmon NFH are shown in Fig. 15d. It can be seen that this NFH has some resemblance to the Airy grating obtained by Minovich et al. [8], however, here the slits are replaced by holographic lobes, allowing for generation of the desired amplitude and phase distribution. In this approach no optimizations are required and both amplitude and phase modulations are encoded within the NFH, however, limiting factors exist for this approach as well. The first relates to the resolution in fabrication of the plasmonic holograms, e.g. the smallest feature that can be fabricated. These are limited by the fabrication process, materials, resist thickness and the electron-beam lithography. The effect of this limitation can be seen in the lobe widths of the generated plasmonic HG beams in Fig. 16. It can be seen that the measured width of the lobes are narrower than those of the simulated beam. This difference resides in the fabrication resolution, which its affect can be is seen in HG(1,0) NFH presented in Fig. 10b. The structural edges of the lobes of the fabricated plasmonic hologram narrow rapidly with many discontinuities, thus affecting the generated HG plasmon.

The second factor resides in the total length of the plasmonic hologram, e.g. the number of periodic cycles within the hologram. As mentioned in Section 5, SPPs are excited over a finite propagation distance. However, for an ideal hologram the field emanating from it should be well defined in a one-dimensional plane along the y-axis, namely $A(y)$, $\phi(y)$. Since free-space holograms are

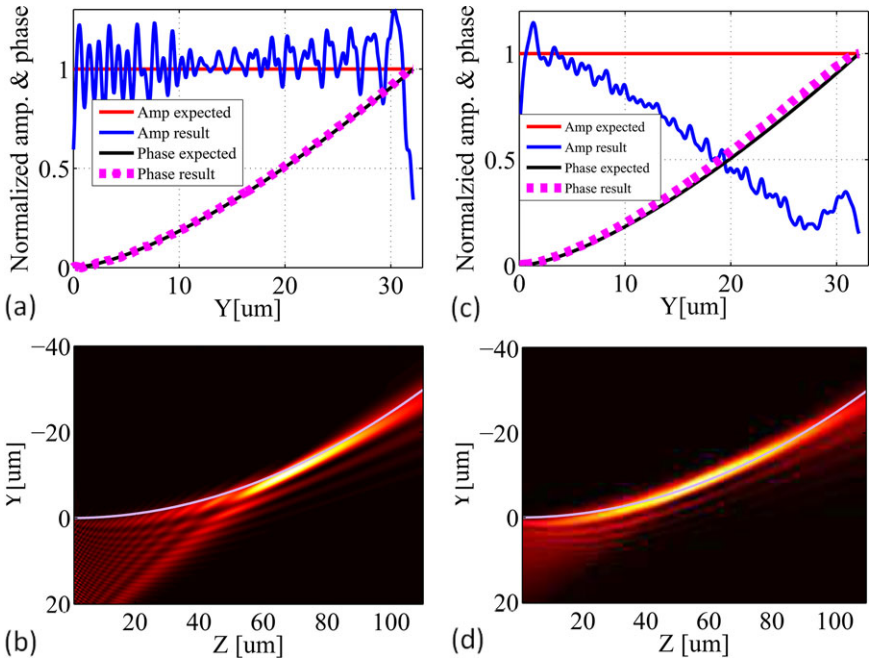


Figure 17 Expected and resulting amplitude (red and blue curves, respectively) and accumulated phase (black and purple dashed curves, respectively) emanating from (a) single-cycle mask and (c) nine-cycles mask, for a parabolic curve. Plasmonic beam intensity emanating from (b) single-cycle, and (d) nine-cycles masks. Reproduced with permission [13].

usually optically thin, this is easily obtained by a planar hologram or SLM. In the case of the plasmons, the finite nature of the hologram along the$z$-axis will affect the generated plasmonic beam. In principle, a single-cycle NFH can be realized, however, its coupling efficiency will be very low owing to the broader spatial frequencies in its Fourier plane. To show this affect, Fig. 17 presents simulations of the resulting amplitude and accumulated phase generated by a single-cycle phase hologram, and a nine-cycle phase hologram. Adding additional cycles results in two different effects – while the coupling efficiency is increased, the resulting amplitude and phase diverge from their target values. It is seen that the amplitude is not constant, as expected from a phase-only hologram, and the accumulated phase increases for larger y values. These changes lead to a deviation from the target trajectory and intensity distribution of the beam (Fig. 17). Our study shows that the optimum number of cycles for the holograms we designed is of about five periods.

The third factor relates to the encoding scheme of the hologram and may be encountered when high spatial frequencies are encoded. Equation (5) assumes a lack of dependency between the transverse and propagation coordinates. However, when highly nonparaxial beams are encoded, this assumption may no longer hold, resulting in a distortion of the beam's wavefront. Furthermore, encoding of high spatial frequencies is accompanied by a smaller geometrical features in the plasmonic hologram, which in turn, relates to the fabrication resolution. It is also noticeable that this approach does not easily allow to change dynamically the properties of the launched plasmonic beam, as demonstrated in [10, 71] since once fabricated, the NFH cannot be changed.

We note here that since the NFH in principle is not limited by any paraxial considerations, nonparaxial self-accelerating beams, such as the half-Bessel, Mathieu and Weber beams [72, 73] can also be realized. These are exact solutions of the nonparaxial Helmholtz equation (Eq. (3)), and therefore enable one to achieve higher rates of acceleration, compared to the paraxial Airy beam. The realization of such nonparaxial SPP beams based on plasmonic NFH have also been demonstrated [14], and the comparison between a paraxial Airy plasmon and a nonparaxial Weber plasmon, having the same trajectory, is presented in Fig. 18.

7.2. Plasmonic cosine-Gauss beam

The cosine-Gauss beam is another kind of "nondiffracting" beam, which propagate along a straight trajectory. It is the two-dimensional equivalent of the three-dimensional Bessel beam, and it is characterized by a transverse infinite cosine field amplitude profile, multiplied by a Gaussian envelope in order to truncate it [74]. It can be mathematically described as the inference between two plane waves, having the same relative angle with respect to the optical axis (Fig. 19a (top)). It has been generated as a nondiffracting plasmonic beam by three different groups by three different methods; one based on a nonholographic method [11], one based on a holographic far-field method [75] and one based on the plasmonic NFH [15]. Its first demonstration by Lin et al. [11] was based on the realization of two slanted metallic grating, in order to launch the two relative angled plane waves, and it was illuminated with a Gaussian beam that acted as the Gaussian envelope. Both measured beam and the structure can be seen in Fig. 19a.

Similarly to the approach used in [10], Xiao et al. [75] used an SLM for generating the cosine-Gauss beam in free-space using a SLM, which then illuminated a metallic

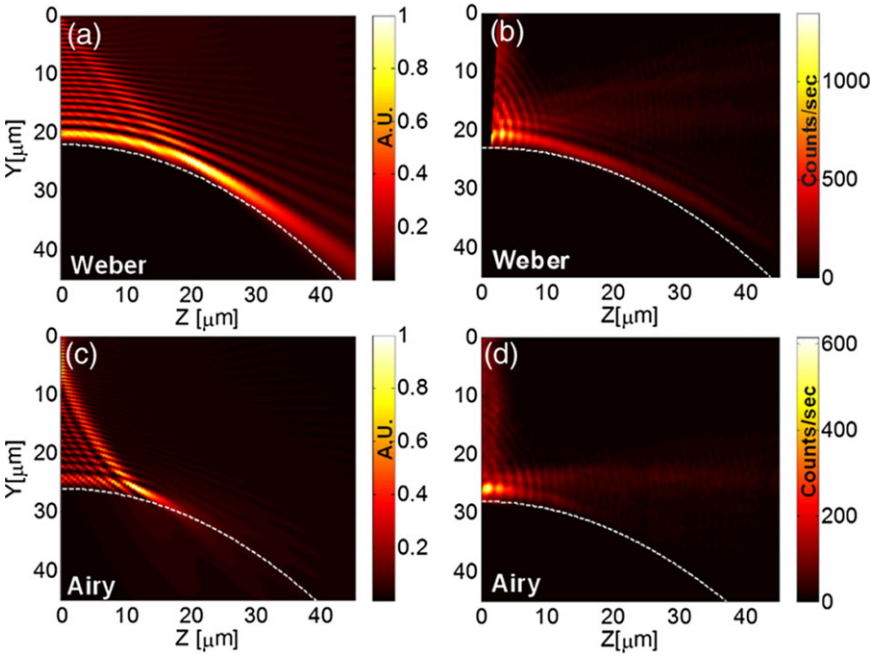


Figure 18 Comparison between measured and simulated Airy and Weber beams. Numerical simulations (a) and (c) and NSOM measurements (b) and (d) of the intensity distribution of the Weber beam and its equivalent Airy beam. The dashed white line represents the analytical trajectory of the beam. Reproduced with permission [14].

grating in order to couple it to a SPP (Fig. 19c). In [15], a plasmonic NFH was encoded with the amplitude and phase of the cosine-Gauss beam, resulting in the hologram presented in Fig. 19b (bottom). The similarity between the NFH obtained from the optical field properties in Fig. 19b and the structure based on the two slanted gratings in Fig. 19a can be clearly seen, together with the generated beams. The main difference between the two is that by using the slanted gratings the amplitude envelope of the beam resides in the illuminating beam, while in the case of the NFH all the information is contained within the hologram and the illuminating beam can be a simple plane wave. Using the illuminating beam in order to define the envelope of the plasmonic beam can be advantageous for dynamically controlling the properties of the launched plasmonic beam, while the NFH can be used for the generation of more complex plasmonic beams. For example, the cosine-Gauss beam can actually be manipulated is such a manner that will allow to compensate for plasmonic losses [76]. This is achieved by manipulating the transverse amplitude of the two plane waves in a manner that allows for giving more energy to the side lobes of the resulting beam. In this way the central lobes of the plasmonic cosine-Gauss beams can maintain a fixed intensity and even a growing intensity, in spite of the plasmonic losses. These kinds of beams require the precise control over the amplitude profile of the beam and this can be provided by the plasmonic NFH. Figure 20 shows the experimental realizations of the regular, fixed intensity and linearly growing plasmonic cosine-Gauss beam.

7.3. Plasmonic 1D Bessel beam

Both the Airy and cosine-Gauss plasmonic beams, previously discussed, are exact solutions of the 2D Helmholtz equation, thus making their generation mostly an experimental challenge. However, the Bessel beam is not a solution of the 2D Helmholtz equation, but of the 3D Helmholtz equation, thus making its generation highly challenging both conceptually and experimentally. In 2011 it was suggested that a 1D Bessel plasmon can actually be realized in a unique structure composed of multiple metal/dielectric layers [77, 78].

A close similarity exists between the cosine-Gauss beam and the 1D Bessel beam [79], which lead to the generation of a 1D Bessel-like plasmon, based on the same methods for generating the cosine-Gauss plasmon [79, 80]. The generation of several Bessel-like plasmonic beams, and the difference between the measured 1D Bessel-like, the calculated 1D Bessel-like, the exact 1D Bessel and the cosine-Gauss beam profiles can be seen in Figs. 21a and b. In 2013, using only a single ridge as the slanted grating, and controlling the relative phases between them, both $J_0(y)$ 1D Bessel-like, and its higher-order $J_1(y)$ Bessel-like beam have also been realized [80] (Fig. 21c). In order to generate the exact 1D Bessel plasmon the holographic approach of the plasmonic NFH has been utilized in 2015 [76], and the measurement of the resulting intensity distribution of the 1D Bessel plasmon can be seen in Fig. 21d.

8. Holographic spectral shaping of surface-plasmons

Apart from fully controlling the plasmonic wavefront, the binary plasmonic NFH can also be utilized for tailoring the energy spectrum of the SPP excitation. In a similar fashion to Section 6.1, the grating manifests an interference between a free-space broadband incident wave as a reference and a spectrally (i.e. temporally) shaped SPP wave as

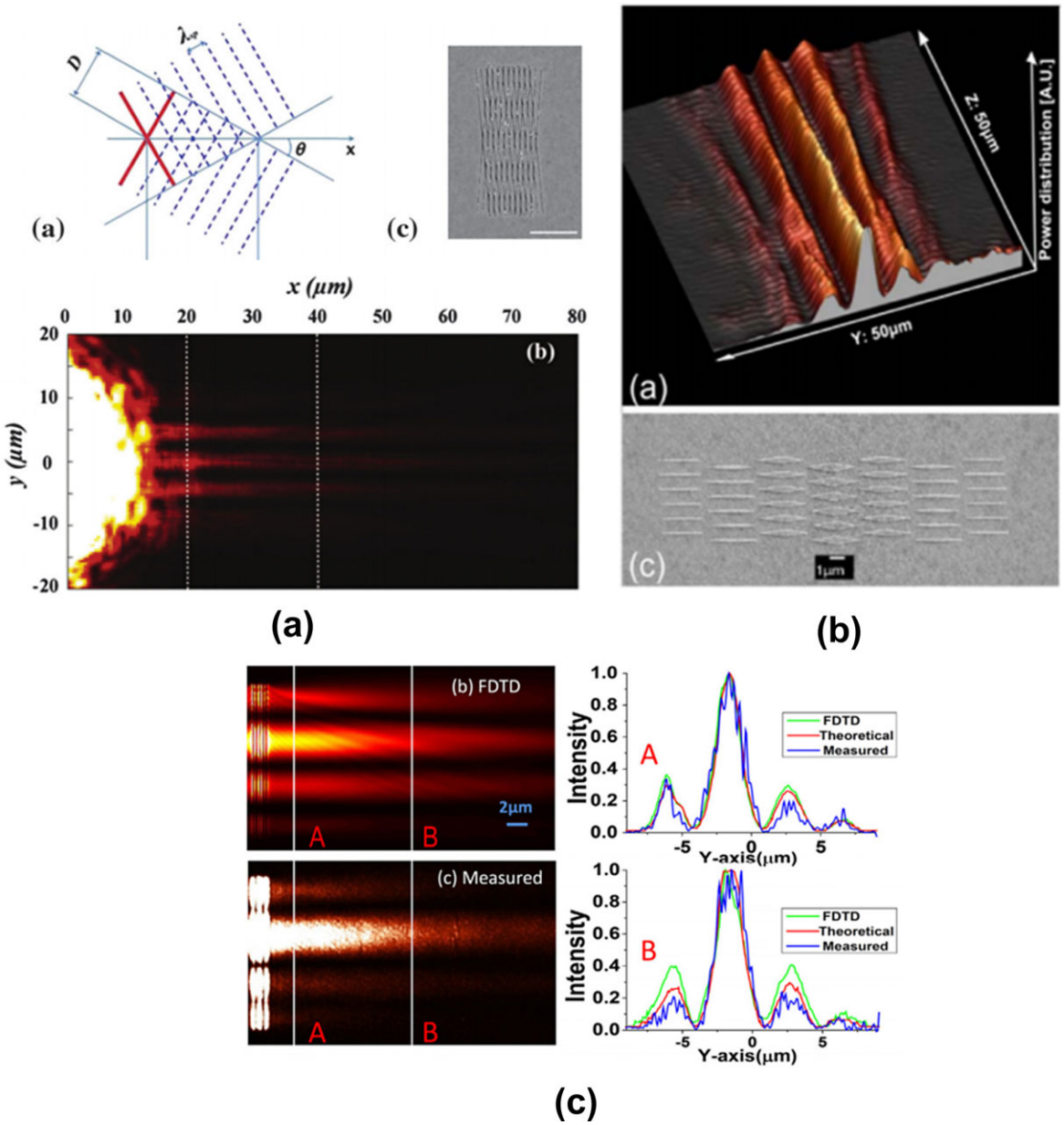


Figure 19 (a) Concept for generating the cosine-Gauss beam according to [11] (top left), the two slanted gratings structure (top right), and the generated beam (bottom). (b) SEM image of the plasmonic NFH realized for the cosine-Gauss beam according to [15] (bottom), and the generated beam from that NFH. (c) Simulated (top left) and measured (bottom left) cosine-Gauss beam and their corresponding cross sections (right) according to [75]. Reproduced with permission [11, 15, 75].

the signal. The binary encoding utilizes the Fourier relation found between the longitudinal thickness modulation pattern $t(z)$ and the wave-number distribution supported by the hologram. An equation similar to Eq. (5) can be derived, but this time the modulation is along the propagation coordinate$z$, whereas earlier in Eq. (5), the modulation was along the transverse coordinate y [16]:

$$t(z) = \frac{h_0}{2}\left(1 + \text{sign}\left\{\cos\left[\frac{2\pi}{\Lambda}z + \Phi(z)\right] - \cos\left[\pi q\,(z)\right]\right\}\right), \tag{8}$$

where h_0 is the ridge height and $q(z)$ is the amplitude modulation of the duty cycle, defined as $q(z) \equiv \sin^{-1}[A(z)]/\pi$. Illuminating this hologram with a broadband, relatively planar beam with a central wavelength obeying Eq. (3), with $k_G = 2\pi/\Lambda$, will launch an SPP wave having the complex-field spectrum represented by the Fourier transform (Fig. 22):

$$U\,(k - k_G) = \int A\,(z)\exp\left[i\,\Phi\,(z)\right]\exp\left(-ikz\right)dz. \tag{9}$$

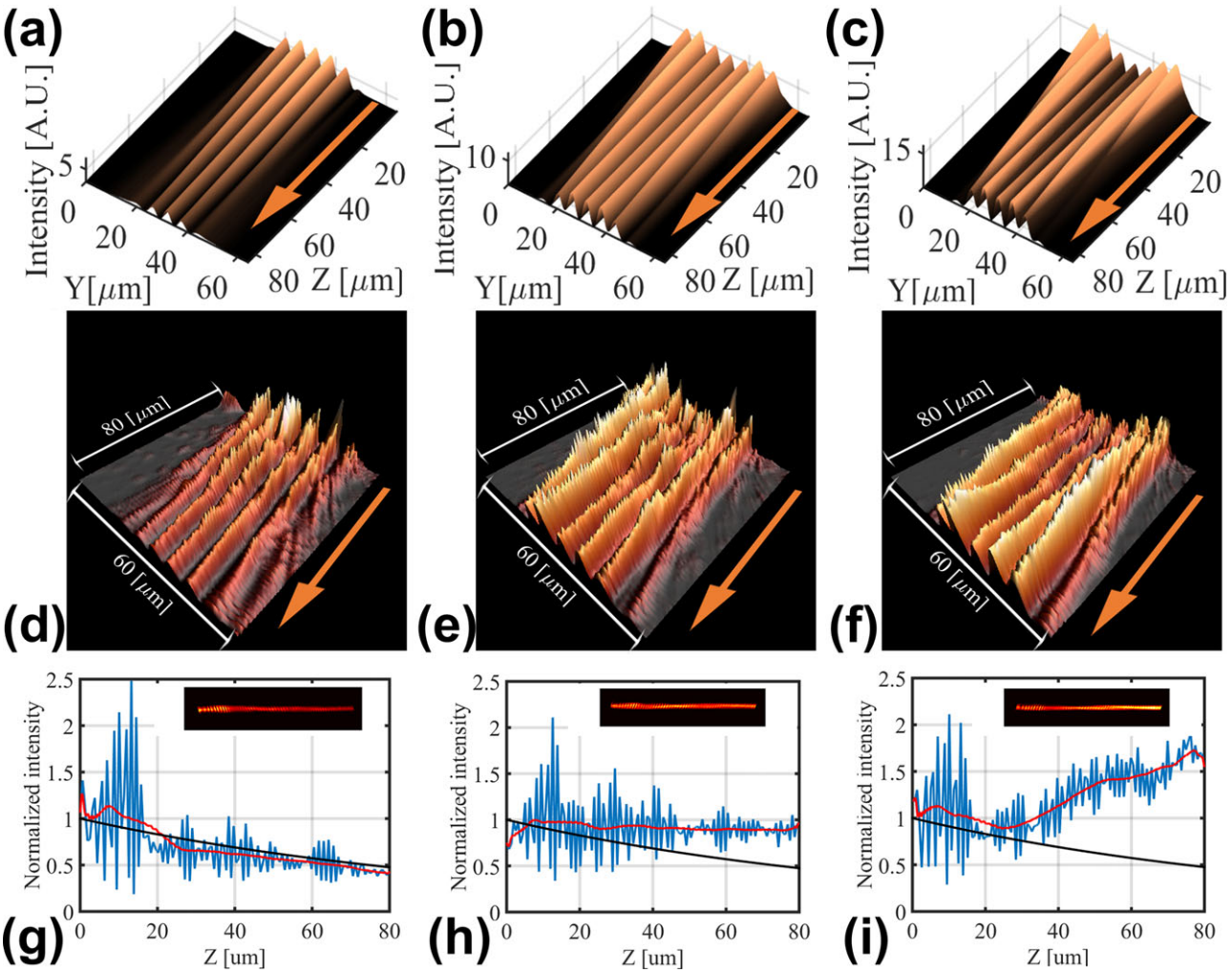


Figure 20 Measurements of (a) a regular, (b) fixed intensity, and (c) linearly growing intensity plasmonic cosine-Gauss beam. (orange arrow represents the direction of propagation). (d)–(f) Averaged intensity behavior exhibited by the central lobe of each beam (blue/red curves), respectively. The Black line represents the theoretical plasmonic decay rate. Reproduced with permission [76].

We note that this spectrum contains not only an amplitude term, for efficiently launching or eliminating certain frequencies, but it also contains a phase term, unlike any previous method. This phase term may open up exciting opportunities for research in interference and temporal shaping of SPP waves. The generation of plasmonic broadband excitation, in the form of a HG(3,0) in its spectrum, is presented in Fig. 22c. We note that simultaneous spectral- and spatial-shaping holograms, having $\Phi(y, z)$ and $q(y, z)$ where z and y are longitudinal and transverse coordinates, respectively, are certainly possible under certain conditions [16].

The ability to control the energy spectrum of the excited SPP could be used for light manipulation on a chip, where ultrashort pulses can couple to SPPs while adding/removing chirp, or splitting a single pulse into several pulses, creating a pulse train.

The encoding of a spectral NFH employs two criteria that ensure proper sampling along both the spectral dimension, represented by k_0 (or $\lambda_0 \equiv 2\pi/k_0$), and the SPP propagation dimension z [81]. The first is a sufficiently large separation of the different grating orders, yielding $\Delta k < 2k_c/3$, where Δk is the bandwidth of the desired spectral response. The second criterion deals with diminishing the effect of the finite longitudinal resolutiondz, which distorts the shape unless $\Delta k \ll 2\pi/dz$.

A different fundamental limitation is represented by the trade-off between the efficiency, or length in the z-axis, and the bandwidth of the shaped SPP. This stems from the Fourier relation between the hologram modulation $t(z)$ (Eq. (6)) and the spectral coupling efficiency $|U(k - k_G)^2|$ (Eq. (7)). As an intuitive example, consider a uniform square grating having N cycles. When N is very large e.g. $N = 1000$, then the excited SPP results from the constructive interference of N scattered waves, resulting in a very high efficiency for a very narrow wave-number spectrum. By contrast, for $N = 1$, this ridge could couple an extremely wide-band SPP, but the

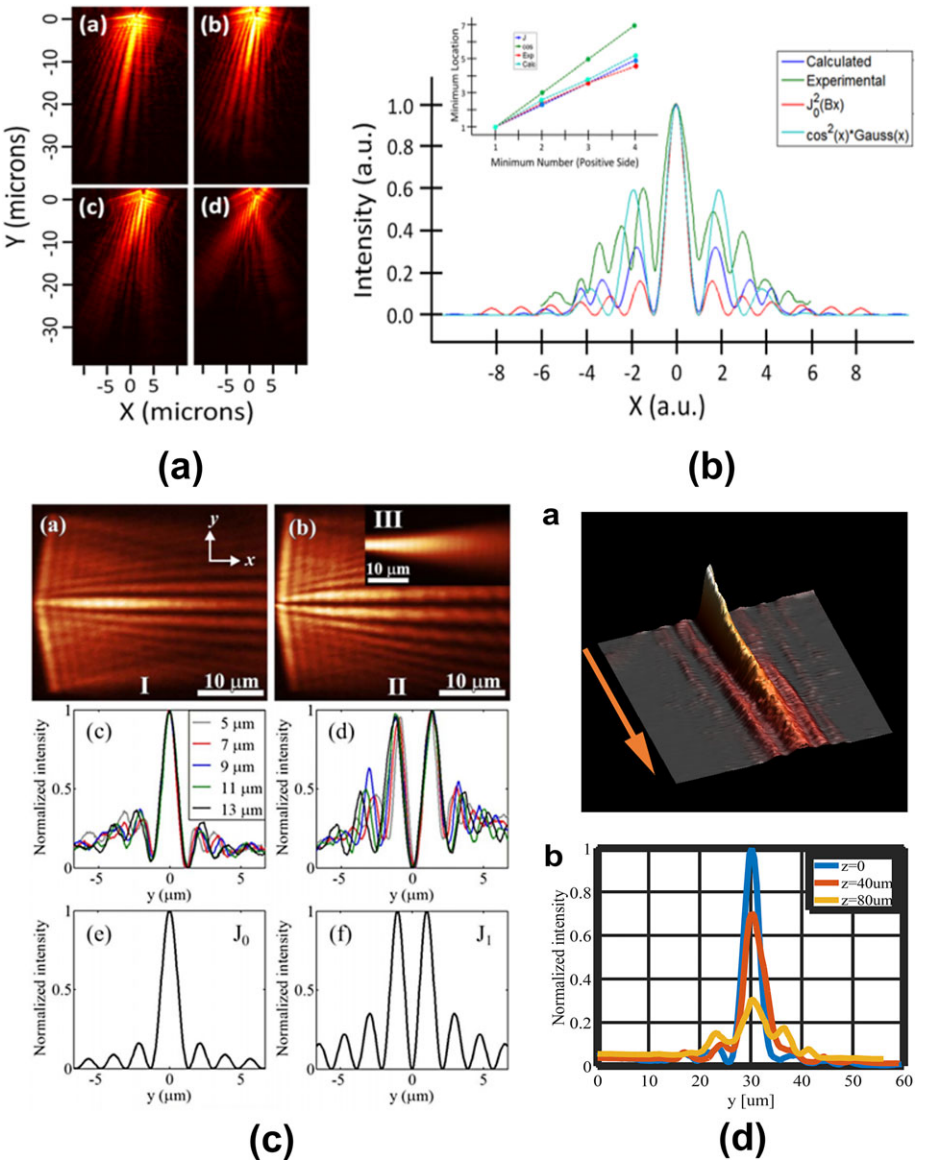


Figure 21 (a) Measurements of several kinds of plasmonic Bessel-like beams. (b) Comparison between the measured 1D Bessel-like, the calculated 1D Bessel-like, the exact 1D Bessel and the cosine-Gauss beam profiles according to [79]. (c) Generation of $J_0(y)$ and $J_1(y)$ 1D Bessel-like plasmonic beams (top). The inset shows a regular Gaussian plasmon, and the bottom shows the relevant measured and theoretical cross sections. (d) Generation of an exact 1D Bessel plasmon using a plasmonic NFH (top), together with three cross-sections along its propagation according to [76]. Reproduced with permission [76, 79, 80].

efficiency is impaired, as no constructive interference takes place.

9. Summary and outlook

In this manuscript we reviewed the history and development of surface-plasmon wavefront and spectral shaping, from the early stages of SPP excitations, to the most recent advanced plasmonic near-field holography. We presented several approaches that have been used to enable the control of the plasmonic wavefront, both nonholographic and holography-based methods, with a focus on binary plasmonic NHF. We showed that these can be used to generate conceptually any arbitrary plasmonic beam and even to compensate for plasmonic losses. We believe that the advancements in this field, providing higher control over the plasmonic-wave properties, are crucial for the realization of new applications in plasmonics, and towards the commercialization of plasmonics in the nanotechnology industry. Moreover, although surface-plasmons are by definition near-field waves, the desire to control the near-field properties of other types of waves has drawn considerable attention in recent years. For example, shaping the nonlinear near-file emitted from nanostructures [82], near-field nonlinear on-axis and volume holography at the second harmonic [83, 84], acoustic near-field holography [85], microwave near-field holography [86], and more [87–89]. The methods and approaches presented in this manuscript may be of interest and applications to these waves, and others, as well.

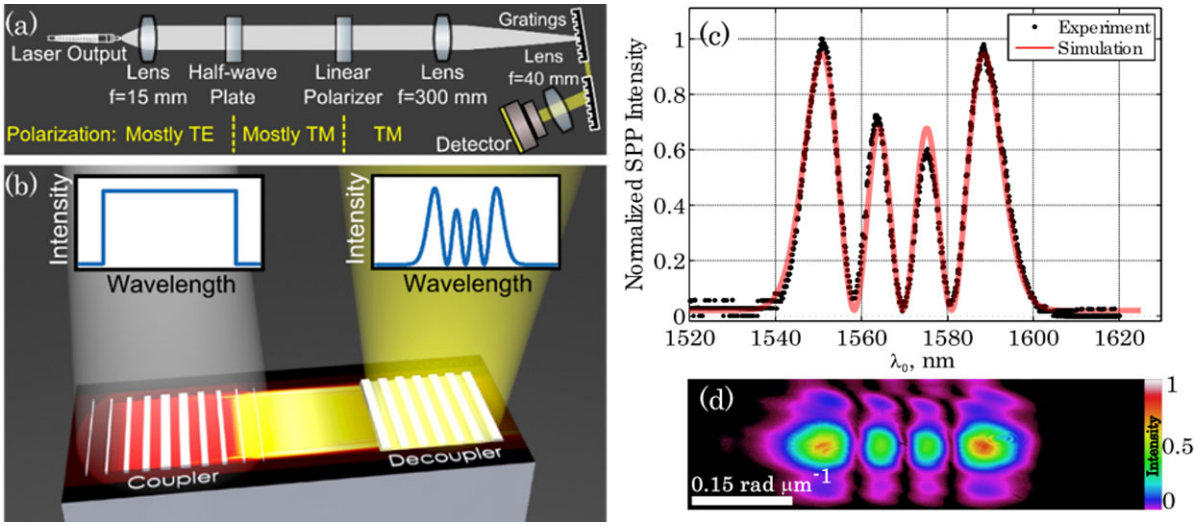


Figure 22 Launching a spectrally tailored plasmonic wave using a holographic grating: (a) the experimental setup and (b) an illustration for a holographic coupling grating (left) designed according to Eq. (8). (c) Plasmonic decoupled spectra for the HG(3,0) shape. (d) The measured far-field light diffraction from the NFH showing the fabricated NFH in k-space. Reproduced with permission [16].

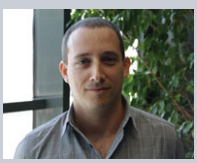
Acknowledgements. This work was supported by the Israel Science Foundation, grant no 1310/13 and by the Marko and Lucie Chaoul Chair in Nano-Photonics.

Received: 30 September 2015, **Revised:** 24 January 2016, **Accepted:** 15 February 2016

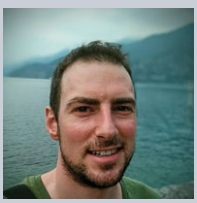
Published online: 18 March 2016

Key words: Surface plasmons, plasmonic holograms, plasmonic beam shaping, surface plasmon holography, near-field holograms, spectral holograms, spectral shaping, plasmonic loss compensation.

Ady Arie received his Ph.D. degree in Electrical Engineering from Tel-Aviv University in 1992. After spending two years as a postdoctoral scholar at Stanford University he joined the School of Electrical Engineering at Tel-Aviv University, where he is at present a Professor of Electrical Engineering and holds the Chaoul Chair in Nano-photonics. His research in the last years is in the areas of nonlinear optics, electron microscopy, plasmonics and hydrodynamics. Prof. Arie is a Fellow of the Optical Society.



Itai Epstein received his B.Sc in physics in 2003 and M.Sc in physics (Magna Cum Laude) in 2007, from Tel Aviv University, Israel. He is currently studying towards a PhD in the Faculty of Electrical Engineering, Department of Physical Electronics, Tel Aviv University. His research focuses on investigating surface plasmon waves as fundamental wave phenomena, and in particular, diffraction from periodical structures, holography, beam shaping, and nonlinear optics with surface-plasmon waves.



Yuval Tsur received his M.Sc. in Physics from the Tel Aviv University in 2013, studying the application of doped silver halide crystals for Mid-IR lasing. He is currently pursuing a Ph.D. in Physics at the Tel-Aviv University, researching the applications of holograms for shaping surface-plasmon, water and electron waves. Yuval is currently part-time employed at Applied Materials Israel Ltd. as an R&D Physicist.

References

[1] S. Maier, Plasmonics: Fundamentals and Applications (Springer, NY, 2007).
[2] W. L. Barnes, A. Dereux, and T. W. Ebbesen, Nature **424**, 824–830 (2003).
[3] N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.P. Tetienne, F. Capasso, and Z. Gaburro, Science **334**, 333 (2011).
[4] I. Epstein, I. Dolev, D. Bar-Lev, and A. Arie, Phys. Rev. B **86**, 205122 (2012).
[5] T. W. Ebbesen, C. Genet, and S. I. Bozhevolnyi, Phys. Today **61**(5), 44 (2008).
[6] D. K. Gramotnev and S. I. Bozhevolnyi, Nature Photon. **4**, 83–91 (2010).
[7] J. N. Anker, W. P. Hall, O. Lyandres, N. C. Shah, J. Zhao, and R. P. Van Duyne, Nature Mater. **7**, 442–453 (2008).
[8] A. Minovich, A. E. Klein, N. Janunts, T. Pertsch, D. N. Neshev, and Y. S. Kivshar, Phys. Rev. Lett. **107**, 116802 (2011).
[9] L. Li, T. Li, S. M. Wang, C. Zhang, and S. N. Zhu, Phys. Rev. Lett. **107**, 126804 (2011).

[10] P. Zhang, S. Wang, Y. Liu, X. Yin, C. Lu, Z. Chen, and X. Zhang, Opt. Lett. **36**, 16, 3191–3193 (2011).
[11] J. Lin, J. Dellinger, P. Genevet, B. Cluzel, F. de Fornel, and F. Capasso, Phys. Rev. Lett. **109**, 093904 (2012).
[12] L. Li, T. Li, S. M. Wang, and S. N. Zhu, Phys. Rev. Lett. **110**, 046807 (2013).
[13] I. Epstein and A. Arie, Phys. Rev. Lett. **112**, 023903 (2014).
[14] A. Libster-Hershko, I. Epstein, and A. Arie, Phys. Rev. Lett. **113**, 123902 (2014)
[15] I. Epstein and A. Arie, J. Opt. Soc. Am. B **31**, 1642–1647 (2014).
[16] Y. Tsur, I. Epstein, and A. Arie, Opt. Lett. **40** (7) 1615 (2015).
[17] N. Yu and F. Capasso, Nature Mater. **13**, 139 (2014).
[18] A. E. Minovich, A. E. Miroshnichenko, A. Y. Bykov, T. V. Murzina, D. N. Neshev, and Y. S. Kivshar, Laser Photon. Rev. **9**, 195 (2015).
[19] Z. Bomzon, V. Kleiner, and E. Hasman, Appl. Phys. Lett. **79**, 1587 (2001).
[20] P. Genevet and F. Capasso, Rep. Prog. Phys. **78**, 24401 (2015).
[21] D. Gabor, Nature **161**, 777–778 (1948).
[22] M. Lehmann and H. Lichte, Microsc. Microanal. **8**, 447–466(2002).
[23] B. R. Brown and A. W. Lohmann, Appl. Opt. **5**, 967–969, 1966.
[24] W. Lee, App. Opt. **21**, 3661–3669 (1979).
[25] J. J. Cowan, Opt. Commun. **12**, 373 (1974).
[26] A. Hohenau, J. R. Krenn, A. L. Stepanov, A. Drezet, H. Ditlbacher, B. Steinberger, A. Leitner, and F. R. Aussenegg, Opt. Lett. **30**, 893 (2005).
[27] E. Devaux, J. Y. Laluet, B. Stein, C. Genet, T. Ebbesen, J.-C. Weeber, and A. Dereux, Opt. Exp. **18**, 20610 (2010).
[28] S. Randhawa, M. U. González, J. Renger, S. Enoch, and R. Quidant, Opt. Exp. **18**, 14496 (2010).
[29] H. Ditlbacher, J. R. Krenn, G. Schider, A. Leitner, and F. R. Aussenegg, Appl. Phys. Lett. **81**, 1762 (2002).
[30] J.-Y. Laluet, E. Devaux, C. Genet, T. W. Ebbesen, J.-C. Weeber, and A. Dereux, Opt. Exp. **15**, 3488 (2007).
[31] J.-S. Bouillard, P. Segovia, W. Dickson, G. A. Wurtz, and A. V. Zayats, Sci. Rep. **4**, 7234 (2014).
[32] G. A. Siviloglou and D. N. Christodoulides, Opt. Lett. **32**, 979–981 (2007).
[33] G. A. Siviloglou, J. Broky, A. Dogariu, and D. N. Christodoulides, Phys. Rev. Lett. **99**, 213901 (2007).
[34] A. Salandrino and D. N. Christodoulides, Opt. Lett. **35**, 12 (2010).
[35] A. E. Minovich, A. E. Klein, D. N. Neshev, T. Pertsch, Y. S. Kivshar, and D. N. Christodoulides, Laser Photon. Rev. **8**, 221–232 (2014).
[36] L. Li, T. Li, S. Wang, S. Zhu, and X. Zhang, Nano Lett. **11**, 4357 (2011).
[37] L. Li, T. Li, S. M. Wang, and S. N. Zhu, Opt. Lett. **37**, 5091 (2012).
[38] B. Gjonaj, J. Aulbach, P. M. Johnson, A. P. Mosk, L. Kuipers, and A. Lagendijk, Nature Photon. **5**, 360–363 (2011).
[39] B. Gjonaj, J. Aulbach, P. M. Johnson, A. P. Mosk, L. Kuipers, and A. Lagendijk, Phys. Rev. Lett. **110**, 266804 (2013).
[40] A. Drezet, D. Koller, A. Hohenau, A. Leitner, F. R. Aussenegg, and J. R. Krenn, Nano Lett. **7**, 1697 (2007).
[41] Y. Liu, S. Palomba, Y. Park, T. Zentgraf, X. Yin, and X. Zhang, Nano Lett. **12**, 4853 (2012).
[42] G. Li and J. Zhang, Sci. Rep. **4**, 5914 (2014).
[43] K. Toma, Y. Masaki, M. Kusaba, K. Hirosawa, and F. Kannari, J. Appl. Phys. **118**, 103102 (2015).
[44] B. Hecht, H. Bielefeldt, L. Novotny, Y. Inouye, and D. Pohl, Phys. Rev. Lett. **77**, 1889 (1996).
[45] M. V. Bashevoy, F. Jonsson, A. V. Krasavin, N. I. Zheludev, Y. Chen, and M. I. Stockman, Nano Lett. **6**, 1113 (2006).
[46] E. Eizner and T. Ellenbogen, Appl. Phys. Lett. **104**, 223301 (2014).
[47] P. E. Landreman and M. L. Brongersma, Nano Lett. **14**, 429 (2014).
[48] P. Biagioni, J.-S. Huang, and B. Hecht, Rep. Prog. Phys. **75**, 024402 (2012).
[49] V. Coello and S. Bozhevolnyi, Opt. Commun. **282**, 3032 (2009).
[50] I. Smolyaninov, D. Mazzoni, J. Mait, and C. Davis, Phys. Rev. B **56**, 1601 (1997).
[51] M. W. Knight, N. K. Grady, R. Bardhan, F. Hao, P. Nordlander, and N. J. Halas, Nano Lett. **7**, 2346 (2007).
[52] D. Chang, A. S. Rensen, P. Hemmer, and M. Lukin, Phys. Rev. Lett. **97**, 053002 (2006).
[53] L. Yin, V. K. Vlasko-Vlasov, J. Pearson, J. M. Hiller, J. Hua, U. Welp, D. E. Brown, and C. W. Kimball, Nano Lett. **5**, 1399 (2005).
[54] O. Avayu, O. Eisenbach, R. Ditcovski, and T. Ellenbogen, Opt. Lett. **39**, 3892 (2014).
[55] J. Lin, J. P. B. Mueller, Q. Wang, G. Yuan, N. Antoniou, X.-C. Yuan, and F. Capasso, Science **340**, 331 (2013).
[56] I. Epstein and A. Arie, Opt. Lett. **39**, 3165 (2014).
[57] O. Avayu, I. Epstein, E. Eizner, and T. Ellenbogen, Opt. Lett. **40**, 1520–1523 (2015).
[58] B. Gallinet and O. J. F. Martin, ACS Nano **5**, 8999 (2011).
[59] Y.-G. Chen, Y. Wang, and Z.-Y. Li, Plasmonics **9**, 1057 (2014).
[60] Y. Chen, Y. Wang, and Z. Li, Opt. Lett. **39**, 339 (2014).
[61] S. I. Bozhevolnyi and B. Vohnsen, Phys. Rev. Lett. **77**, 3351–3354 (1996)
[62] M. Ozaki, J. Kato, and S. Kawata, Science **332**, 23908–23920 (2011).
[63] Y Chen, J. Fu, and Z. Li, Opt. Exp. **19** (2011).
[64] Y. Chen, M. Zhang, L. Gan, X. Wu, L. Sun, J. Liu, J. Wang, and Z. Li, Opt. Exp. **21**, 17558 (2013).
[65] Y. Chen, L. Huang, L. Gan, and Z. Li, Light: Sci. Applic. **1**, e26 (2012).
[66] I. Dolev, I. Epstein, and A. Arie, Phys. Rev. Lett. **109**, 203903 (2012).
[67] P. Genevet, J. Lin, M. A. Kats, and F. Capasso, Nature Commun. **3**, 1278 (2012).
[68] E. Greenfield, M. Segev, W. Walasik, and O. Raz, Phys. Rev. Lett. **106**, 213902 (2011).
[69] L. Froehly, F. Courvoisier, A. Mathis, M. Jacquot, L. Furfaro, R. Giust, P. A. Lacourt, and J. M. Dudley, Opt. Exp. **19**, 16455 (2011).
[70] B. Gallinet, J. Butet, and O.J. F. Martin, Laser Photon. Rev. **9**, 577–603, (2015)
[71] D. Bar-Lev, A. Arie, J. Scheuer, and I. Epstein, J. Opt. Soc. Am. B **32**, 923–932 (2015).

[72] I. Kaminer, R. Bekenstein, J. Nemirovsky, and M. Segev, Phys. Rev. Lett. **108**, 163901 (2012).
[73] P. Zhang, Y. Hu, T. Li, D. Cannan, X. Yin, R. Morandotti, Z. Chen, and X. Zhang, Phys. Rev. Lett. **109**, 193901 (2012).
[74] J. C. Gutiérrez-Vega and M. A. Bandres, J. Opt. Soc. Am. A **22**, 289 (2005).
[75] K. Xiao, S. Wei, C. Min, G. Yuan, S. W. Zhu, T. Lei, and X.-C. Yuan, Opt. Exp. **22**, 13541 (2014).
[76] I. Epstein, R. Remez, Y. Tsur, and A. Arie, Optica **3**, 15–19 (2016).
[77] C. J. Zapata-Rodríguez, S. Vuković, M. R. Belić, D. Pastor, and J. J. Miret, Opt. Exp. **19**, 19572 (2011).
[78] C. J. Zapata-Rodríguez, D. Pastor, V. Camps, M. T. Caballero, and J. J. Miret, J. Nanophoton. **5**, 051801 (2011).
[79] C. J. Regan, L. Grave de Peralta, and A. A. Bernussi, J. Appl. Phys. **112**, 103107 (2012).
[80] C. Garcia-Ortiz, V. Coello, Z. Han, and S. Bozhevolnyi, Opt. Lett. **38**, 905 (2013).
[81] R. Shiloh and A. Arie, Opt. Lett. **37**, 3591 (2012).
[82] D. Wolf, T. Schumacher, and M. Lippitz, Nature Comm. **7**, 10361 (2015).
[83] S. Trajtenberg-Mills, I. Juwiler, and A. Arie, Laser Photon. Rev. **9**, 40–44, (2015).
[84] X.-H. Hong, B. Yang, C. Zhang, Y.-Q. Qin, and Y.-Y. Zhu, Phys. Rev. Lett. **113**, 163902 (2014).
[85] J. D. Maynard, E. G. Williams, and Y. Lee, J. Acoust. Soc. Am. **78**, 1395 (1985).
[86] R. K. Amineh, A. Khalatpour, H. Xu, Y. Baskharoun, and N. K. Nikolova, Int. J. Biomed. Imag. **2012**, 291494 (2012).
[87] M. Stockmar, P. Cloetens, I. Zanette, B. Enders, M. Dierolf, F. Pfeiffer, and P. Thibault, Sci. Rep. **3**, 1927 (2013).
[88] H. Mühlenbernd, P. Georgi, N. Pholchai, L. Huang, G. Li, S. Zhang, and T. Zentgraf, ACS Photon. **3**, 124–129 (2016).
[89] H. D. Kjeldsena, M. Kaiserb, and M. A. Whittingtonc, J. Neurosci. Methods **253**, 1–9 (2015).