This tutorial aims to provide an extensive overview of methods of generating and shaping light at new frequencies by using nonlinear metasurfaces. We first review methods of manipulating light by using linear metasurfaces, on the basis of local control of the amplitude and phase of transmitted and reflected light. To extend these principles to nonlinear metasurfaces, we first introduce the mechanisms and principles underlying nonlinear interactions in metasurfaces. We then show how to use these principles to control the phase, amplitude, and polarization of emitted nonlinear radiation and how, through careful spatial arrangement of single nonlinear elements on a metasurface, it is possible to tailor the shape of the light emitted through nonlinear interaction. © 2018 Optical Society of America

OCIS codes: (160.3918) Metamaterials; (160.4330) Nonlinear optical materials; (190.4350) Nonlinear optics at surfaces; (190.4360) Nonlinear optics, devices; (230.4320) Nonlinear optical devices; (310.6628) Subwavelength structures, nanostructures

https://doi.org/10.1364/AOP.10.000309
4. Control Over Local Nonlinear Tensor Elements for the Nonlinear Shaping of Light
   4.1. Control Over Amplitude
   4.2. Control Over Phase
   4.3. Control Over Polarization
5. Nonlinear Diffraction—Exploiting the Collective Nonlinear Response
   5.1. Collective Effects on Nonlinear Metasurfaces
   5.2. Nonlinear Photonic Crystals
6. Nonlinear Light Shaping into Beams and Arbitrary Images
   6.1. Nonlinear Focusing of Light
   6.2. Nonlinear Image Encoding on a Metasurface
   6.3. Nonlinear CGH Metasurfaces for Far-Field Beam Shaping
   6.4. Perfect Nonlinear Beam Shaping
   6.5. Arbitrary Shaped Nonlinear Holograms
7. Conclusion and Outlook
Funding
Acknowledgment
References
1. INTRODUCTION

Shaping light, i.e., tailoring the spatial shape of the light’s wavefront, is an essential function of many optical systems. The ability to manipulate the shape of light, which defines some of its most important physical properties, such as the momentum and information it carries as it propagates, enables the distribution of its electro-magnetic (EM) energy into specific spatial locations and control over the properties of light–matter interactions. Therefore, the ability to form optical wavefronts on demand is very appealing for both applications and fundamental science.

The most common and useful description of the spatial evolution of a monochromatic light beam propagating in a well-defined direction is provided by the paraxial approximation of the Helmholtz equation:

$$\partial_z^2 U + \partial_x^2 U + 2ik\partial_z U = 0,$$

where $z$ is the propagation direction, $U$ is the complex envelope of the electromagnetic wave, and $k$ is the wave vector. The eigensolutions to this equation are beams of light that maintain their spatial form while propagating.

The most common beam-type solution to the paraxial Helmholtz equation is a Gaussian beam [1,2]. More general beam types include Hermite–Gauss, Laguerre–Gauss, and Bessel–Gauss beams [1–3]. In addition, there are many other types of beams that serve as solutions to the paraxial wave equation. Each of these solutions has various valuable properties; e.g., Laguerre–Gauss beams carry orbital angular momentum, Bessel beams are nondiffracting [4], and Airy beams move along curved trajectories [5]. Different beam shapes are used to carry energy in various profiles that are tailored for specific applications, e.g., stimulated emission-depletion microscopy [6,7], optical tweezing [8,9], and the optimization of optical gain [10,11]. In addition, owing to the orthonormal nature of the different modes of a given basis, multiple different beams can be used for data multiplexing [12,13].

Shaping light into well-defined beams involves only the basic solutions to the paraxial wave equation. Light can also be shaped into more flexible forms to create arbitrary images in the optical far field. Because the diffraction pattern that appears in the far field is related to the generated beam wavefront through a Fourier transform relation [14], the ability to design arbitrary far-field wavefronts relies on the ability to imprint any arbitrary pattern of amplitude, phase, and polarization on the transverse plane of the beam. To find the patterns required for a specific image, holography methods are often used [15–19], in which the holograms defining the desired diffraction patterns can be calculated analytically, in simple cases, or numerically, for more complex patterns. These numerical calculations are based on computational methods such as fast Fourier transformation, the Gerchberg–Saxton algorithm, and the FiDOC algorithm for phase retrieval [20–25]. Holograms designed by using these techniques are known as computer-generated holograms (CGHs).
Over the past couple of decades, there have been immense advancements in the ability to shape light using two-dimensional optical materials called metasurfaces. These are optically thin structured materials made from subwavelength nanostructures, also referred to as meta-atoms, which allow control over the properties of the light interacting with them. Metasurfaces enable the manipulation of waves in the optical far field and can be used as optical elements for beam shaping and holography. Moreover, metasurfaces also enable the control of light in the intermediate and near-field zones. By tuning the behavior of waves in the intermediate regime, also known as the Fresnel zone, it is possible, for example, to focus beams or engineer desired point spread functions. At shorter distances, at which the behavior of light is governed by its near-field characteristics, judicious manipulation of the energy distribution can be beneficial for various processes, such as nonlinear generation of new frequencies, surface-enhanced Raman scattering, enhancement of Purcell factors, and enhancement of the chirality of molecules.

Whereas linear optics and the shaping of light have been studied since ancient times, nonlinear optics was introduced only half a century ago, and nonlinear beam shaping has been practiced for only the past few years [26–30]. Moreover, most of the research and applications developed over the past decade in the field of optical metasurfaces have focused on the linear shaping of optical wavefronts. Nonlinear optics permits the generation of light at new frequencies and the control of photon–photon interactions by using nonlinear optical materials [31]. Therefore, leveraging light shaping techniques to also control nonlinear light–matter interactions permits the efficient creation of light at new frequencies in predefined shapes. In contrast to the linear shaping of light, in which the wavefront is manipulated by changing the spatial amplitude and phase profiles of existing light, in the nonlinear case, light is directly generated at new frequencies with the desired shape.

To shape light via a nonlinear optical process without necessarily shaping the incident light, there is a need to exert spatial control over the nonlinear susceptibility tensor. However, when natural nonlinear crystals are used, some limitations apply. The value of the nonlinear tensor is a naturally given constant that cannot be changed. It is possible to manipulate only the sign of the quadratic nonlinear tensor, at very limited resolutions [32,33]. In addition, it is necessary to consider the phase mismatch between propagating waves to achieve efficient interaction between them. This phase mismatch reduces the efficiency of the interaction and distorts the shape of the nonlinear output. Recently, a new approach for the nonlinear shaping of light by using metasurfaces has been presented. Owing to the unique properties of metasurfaces, many of the problems that arise when using natural nonlinear materials for nonlinear beam shaping can be alleviated. In recent years, it has been shown that metasurfaces enable efficient local control over the amplitude, phase, and polarization of transmitted and reflected light. Over the past decade, these advancements have led to many applications and discoveries, from optical spin-orbit coupling [34] to the generalized Snell’s law [35]. In parallel, researchers have begun to use these new methods of nanoscale optical design to enable 2D functionalities that allow the operation of conventional optical elements to be mimicked in thin 2D surfaces and even expand the functionality of the systems beyond what can be achieved through conventional means [36,37]. These concepts of functional metasurfaces for controlling the shape of the light were recently adapted to nonlinear metasurfaces. Such functional nonlinear metasurfaces benefit from advantages such as local and absolute control of the nonlinear tensor, resonant interaction, and phase-mismatch-free interactions, which cannot be achieved in conventional functional nonlinear devices.

Although the linear response of such materials can be explained by effective medium theories and Mie scattering and its nonlocal extensions [38–42], the intriguing optical
nonlinearity of such metasurfaces is still not fully understood. In addition to strong
efforts to study the fundamental nonlinear behavior of metasurfaces [43–47], in
the past few years, nonlinear metasurfaces have been leveraged to achieve unprecedented
functionalities, including the nonlinear shaping of emitted light beams [48–54].

In this tutorial, we briefly outline the background of linear and nonlinear interactions
with optical metasurfaces and general beam shaping concepts. In addition, we present
a broad overview of the linear and nonlinear beam shaping of light by using meta-
surfaces. We mainly focus on linear and nonlinear metasurfaces of the plasmonic type.
Although the field of dielectric nonlinear metasurfaces may provide advantages in
terms of efficiency and bandwidth, it is still evolving. In addition, the underlying
mechanisms that provide the nonlinearity in dielectric metasurfaces are essentially
different. Nevertheless, many of the concepts that are being used for nonlinear shaping
of light by plasmonic metasurfaces will be relevant also for nonlinear shaping of light
by dielectric nonlinear metasurfaces. Therefore, along the manuscript we will shortly
mention some of the progress that was done in this field.

The paper is organized as follows: In Section 2, we begin with a short overview of the
interactions of light with subwavelength structures and continue by presenting the
manipulation of light by using linear metasurfaces. This section is brought as a short
introduction to the readers who are not familiar with the subject of shaping light by
linear metasurfaces. For further reading on this subject we advise the readers to turn to
more comprehensive reviews [55–67]. Readers who are already familiar with this
subject can skip to the next section. The concepts presented in Section 2 in their linear
context are expanded in the following sections to their nonlinear parallels. In Section 3,
we introduce the mechanism of nonlinear interactions in metasurfaces, and in Section 4,
we describe the methods for achieving full control over the nonlinear tensor by using
metasurfaces. Later, in Section 5, we discuss the underlying mechanisms of collective
nonlinear effects in nonlinear metasurfaces. In Section 6, we extend these concepts to
the nonlinear shaping of light and review several recent studies in this field. We also
discuss the possibility of using nonlinear holography for the nonlinear generation of
light with any arbitrary shape. Finally, in Section 7, we summarize and discuss the future
outlook for the nonlinear shaping of light with metasurfaces.

We believe that this tutorial should be helpful for both experts and beginners in the
field of metasurface-based nonlinear beam shaping and that it may serve as a practical
guide for developing new and exciting metasurface devices for optical imaging, spec-
troscopy, communication, and computing.

2. MANIPULATION OF LIGHT WITH LINEAR METASURFACES

Metasurfaces can be used to efficiently and locally control the amplitude, phase, and
polarization of transmitted and reflected light. Therefore, they offer the possibility of
miniaturizing traditional optical elements and realizing new functionalities. Over the
past two decades, a wide variety of optical elements have already been realized by
using metasurfaces, including wave plates [68,69], polarization switches, holograms,
diffractive gratings [70,71], wavelength-selective surfaces [72], and lenses [73–75].
Numerous types of metasurfaces have been studied, spanning from metasurfaces com-
posed of either metallic or high-index dielectric planar assemblies of nanoscale optical
resonators to metasurfaces with designs based on holes in metallic or dielectric films.
Their optical response depends on the individual nanoresonator material used, its mor-
phology, the geometry of the array and the surrounding media. The purpose of this
section is to lay the foundations for shaping light with nonlinear metasurfaces by
understanding the basic principles of linear metasurfaces and ways to apply them
for shaping of light. For further extensive reading on linear metasurfaces, additional comprehensive reviews are available [55–67].

Because the net response of a metasurface stems, most often, from the scattering of light by individual nanoresonators, we will briefly describe the underlying physical mechanism of such scattering. We will mainly focus on the case in which the nanoresonators are metallic nanoparticles and thus support resonances associated with free electron motion, called localized surface plasmon resonances (LSPRs). These resonances lead to enhancement and localization of the EM field in the vicinity of a metallic nanoparticle along with enhanced scattering cross sections. Additionally, LSPRs may potentially give rise to enhanced nonlinear light–matter interactions, thus making them attractive for nonlinear processes, as will be discussed in the following section. In addition to plasmonic resonances, we will briefly discuss the mechanisms and features of the dielectric resonances of nanoparticles as well as their potential to be incorporated into linear and nonlinear beam shaping devices.

2.1. Local Manipulation of Light Scattering by Means of Single and Coupled Nanoresonators

The physics underlying LSPRs is the collective oscillations of free electrons associated with the LSPRs, which can be described by using a driven harmonic oscillator model [56]. In this model, a restoring force is exerted by the positive ions in the metal, which attract the displaced electrons. The phase of the electron cloud motion relative to the phase of the electrical driving force determines the relative phase and amplitude of the scattered light. At resonance, the motion of the electrons lags with respect to the electrical driving force by a phase of $\pi/2$, whereas the full spectral width extends from in-phase motion, for frequencies lower than the resonant frequency, to $\pi$-phase-shifted motion, for frequencies higher than the resonant frequency (Fig. 1). Consequently, when only a single plasmonic resonance is involved, the phase of the scattered light can be controlled only within a range of $\pi$, and such manipulation is accompanied by changes in the scattering amplitude.

![Figure 1](image)

The amplitude and phase of the free electron cloud motion versus the wavelength. Inset: a sketch of a plasmonic nanorod whose electron cloud has been displaced. The displacement behavior of the free electrons at a nanostructure caused by an oscillating EM field can be treated as a driven damped harmonic oscillator. In the vicinity of the resonance wavelength, the phase of the electron motion relative to the driven field changes significantly.
The above description depicts the mechanism of plasmonic resonances only qualitatively; however, one must often consider the exact spectral positions of these resonances to design metasurfaces with the required functionalities. To achieve this goal, one must either perform a rigorous analysis to solve the exact electromagnetic problem or perform numerical calculations using electromagnetic simulation tools or approximate methods. An exact solution to the scattering problem exists only for spheroids [41] and ellipsoids [42], whereas in the general case, the polarizability of a nanoparticle, denoted by \( \alpha \), can be found by using the quasi-static approximation. This approximation assumes that the phase of the applied field is constant over the entire particle volume; therefore, the particle size \( d \) must be much smaller than the wavelength in the surrounding medium, i.e., \( d \ll \lambda_{med} \). For larger particles, dynamical corrections to this approximation are often used. An important result of the quasi-static approximation is the polarizabilities along the principal axes for an ellipsoid with semiaxes \( a_1, a_2, \) and \( a_3 \) and volume \( V = \frac{4\pi}{3}a_1a_2a_3 \) [76]:

\[
\alpha_i = V \frac{\epsilon(\omega) - \epsilon_{med}}{\epsilon_{med} + L_i(\epsilon(\omega) - \epsilon_{med})},
\]

where \( \epsilon(\omega) \) and \( \epsilon_{med} \) are the dielectric constants of the particle and the surrounding medium, respectively, and the \( L_i \), which satisfy \( L_1 + L_2 + L_3 = 1 \), are geometric factors given by:

\[
L_i = \frac{a_i a_j a_k}{2} \frac{d q}{(a_i^2 + q^2)(a_j^2 + q^2)(a_k^2 + q^2)}
\]

where \( f(q) = \sqrt{(q + a_1^2)(q + a_2^2)(q + a_3^2)} \). For a qualitative discussion of the outcomes of the polarizability expression and the conditions under which its denominator vanishes and a resonance condition is fulfilled, one can consider the simplified dielectric constant of a Drude metal below its plasma frequency \( \omega_p \), with a collision angular frequency \( \gamma \), \( \epsilon_{Drude}(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i \gamma \omega} \). The simple Drude model cannot describe interband transition effects, which typically occur in the visible regime for noble metals. Although the above polarizability result was formulated for small particles in the quasi-static regime, it can typically be applied for particles with dimensions of the order of tens of nanometers [77,78]. Moreover, it captures several of the most important features of plasmonic resonances, as described below. First, the different components of the polarizability described in Eq. (2) correspond to spectrally separated resonances, which depend on the semiaxis lengths. Second, in an elongated plasmonic nanoparticle, the resonance along the direction of elongation redshifts as the corresponding semiaxis length increases. Third, the plasmonic resonances redshift as the dielectric constant of the surrounding medium increases. The polarizability of an ellipsoid, with adequate corrections to account for dynamic depolarization and radiative damping [79], can be used to calculate the spectral positions and line shapes of the resonances for various geometries, such as nanorods, nanoplates, and nanospheres.

For larger nanostructures, it is also possible to regard a nanostructure as an effective Fabry–Perot cavity for surface plasmons. This treatment is justified when the characteristic dimension of the structure is comparable to or larger than half the effective surface plasmon wavelength, and it permits the determination of multiple resonance frequencies. For example, in a nanorod, the resonant frequencies can be derived by requiring one round trip of the guided mode to result in a phase accumulation of an integer multiple of \( 2\pi \), as follows:

\[
m\lambda_{eff} = 2L + 2\phi_r,
\]

where \( L \) is the length of the nanorod, \( \phi_r \) is a reflection phase that can be related to the extension of the field outside the ends of the rod, and \( \lambda_{eff} \) is the effective plasmon wavelength, which has been shown to obey the following empirical relation [80]:

\[
\frac{\omega}{\epsilon_{eff}} = \frac{\omega_p}{\epsilon_{Drude}} + \frac{\omega_c}{\epsilon_{Drude}}
\]
where \( n_1 \) and \( n_2 \) are constants related to the geometry and the dielectric environment, respectively, and \( \lambda_p \) is the plasma wavelength. The conditions for LSPRs now become dependent on the effective wavelength \( \lambda_{\text{eff}} \) rather than the free-space wavelength. Typical ratios of \( \lambda / \lambda_{\text{eff}} \) are approximately 2–3 in the visible and near-infrared regimes, thus permitting deep subwavelength sizes for the nanoresonators. Using this model, it is possible to gain valuable insight into the antenna responses of different metallic shapes, such as nanodisks [81–84], nanorods [80,85,86], split-ring resonators (SRRs) [87,88], and V-shaped antennas [35,89,90].

The case of a single plasmonic resonance, which provides only partial control over the phase and amplitude of the scattered light, has been considered above. To achieve complete control over the wavefront, however, phase modulation over the full \( 2\pi \) range is necessary. This can be achieved by considering more complex configurations, such as a single antenna that supports multiple independent resonances [35,91], coupled antenna resonances [92], and array geometries that take advantage of the polarization of light to achieve any desired phase [93,94]. Next, we will discuss coupled plasmonic nanoresonators and the possibility of multiple simultaneous resonances in dielectric nanoresonators. Then we will review the concepts of detour and geometrical phase acquisition.

When two or more metallic nanoparticles are brought into proximity, the near-field and radiative interactions become substantial, and as a result, changes in the optical response of the coupled plasmonic system are evident. Because each of the nanoparticles can be described as a harmonic oscillator, the interaction among closely packed nanostructures can be described by using the coupled harmonic oscillators model [80]. Among the various existing methods, a more thorough description of the changes in the collective modes of such assembled resonators is given by the

\[
\lambda_{\text{eff}} = n_1 + n_2 \left( \frac{\lambda}{\lambda_p} \right),
\]  

(4)

Figure 2

Plasmon hybridization of metallic nanoparticles. Two identical spherical metallic nanoparticles hybridize, thus producing four spectrally separated modes of the interacting system. The charge distributions over the two particles are shown for the different possible modes of the coupled and uncoupled states. The depicted mode splitting can be derived by using either the plasmon hybridization model [92,95,96] or the coupled dipole approximation [77,97,98]. Reprinted with permission from Myroshnychenko et al., Chem. Soc. Rev. 37, 1792–1805 (2008) [77].
plasmon mode hybridization method [92,95,96] or the coupled dipole approximation [97,98]. One of the prominent features of a strongly coupled system that is predicted by these methods is mode splitting, in which the strong near-field interaction results in rapid energy transfer between the individual resonators results in new plasmonic modes of the combined system. An example of this phenomenon is the coupling of the dipolar plasmonic modes of two neighboring metallic spheres, as shown in Fig. 2. The fundamental modes of the two nanoparticles hybridize and produce spectrally separated modes of the interacting system. The aforementioned methods and other analytical techniques can be used to engineer coupled plasmonic nanoparticle geometries that provide greater control over the amplitude and phase of the scattered light. This capability is beneficial for the design of any arbitrary desired wavefront.

In addition to plasmonic metasurfaces, in recent years, several groups have also studied dielectric metasurfaces [70,91,99–103], which are planar devices composed of high-refractive-index dielectric materials. Using dielectric rather than metallic nanostructures offers the possibility to overcome the high absorption and backward scattering amplitudes associated with metals, thereby enabling the realization of nano-devices with high transmission efficiencies. It is important to mention that the physical mechanism of such devices is cardinaly different, as will be described briefly. In the context of beam shaping, it has been shown that by using dielectric nanoparticles, the phase of scattered light can be controlled either by operating near a resonance [91] or by means of continuous phase accumulation through nanoparticles that act as Fabry–Perot resonators [99]. The first method relies on dielectric Mie resonances based on displacement currents. These resonances can either be electrical or magnetic in nature; both types have comparable scattering amplitudes, and can be structured to coexist in the same spectral range, thus giving rise to intriguing interference effects, such as unidirectional scattering and full phase control of the scattered light. More specifically, because each dipolar resonance is associated with a 0 to $\pi$ phase shift of the scattered light, a combination of the two types of resonances, i.e., electrical and magnetic dipole resonances, enables control over the full $2\pi$ range. The second approach relies on treating the high-refractive-index nanostructures as low-Q-factor Fabry–Perot resonators, as also presented in Subsection 2.4 [99]. The effective refractive index for each mode in the resonator is proportional to its diameter, i.e., its dimension parallel to the surface; thus, different diameters result in different phase accumulation. Through this method, it is also possible to control the phase of the scattered light separately for each polarization using edge-truncated elliptical nanoparticles.

The mechanisms described above concern the use of single or coupled particles to control scattered light. However, the spatial arrangement of a metasurface composed of multiple nanoresonators enables further control of the scattered light’s phase pattern, by means of either a location-based detour phase [104,105] or the ability to take advantage of the polarization of the light through what is known as a geometrical phase [106,107].

### 2.2. Detour Phase

Beyond the possibility of controlling the phase of waves scattered from a metasurface by adjusting the response of the inclusions, there is an inherent degree of freedom hidden in the spatial arrangement of the inclusions. This concept, known as the detour phase, is the basis for the design of diffractive optical elements and serves as the underlying approach for the generation of binary and regular CGHs [104,105]. The simplest and best-known example taking advantage of this characteristic of wave interaction is a binary diffraction grating. In this case, without any control over the
phase of the waves scattered from different positions along the grating, the light is deflected into noncollinear directions. In this case, light that is simultaneously scattered from two different locations separated by a distance $\Delta x$ will acquire a relative phase of

$$\Delta \phi = \Delta x \frac{2\pi}{\lambda} \sin(\theta),$$  \hspace{1cm} (5)$$

where $\theta$ is the angle of the scattered light relative to the normal to the plane of the scatterers. The minimum phase difference that must be acquired such that the two beams constructively interfere in the far field is $2\pi$, and thus, constructive interference occurs only for

$$\Delta x \cdot \sin(\theta) > \lambda.$$  \hspace{1cm} (6)$$

This result implies that by designing a metasurface with different zones satisfying Eq. (6) and adjusting its phase profile in accordance with Eq. (5), free manipulation of light beams can be achieved. Figure 3 illustrates the ability to focus light by using the detour phase concept. This concept has been used in combination with metasurfaces to realize various optical elements, including chromatically corrected metasurface lenses [108] and broadband dielectric and plasmonic holograms [71,109].

**Figure 3**

Detour phase. The wavefront can be controlled by means of the transverse spatial arrangement of the transparent and opaque areas in accordance with the acquired detour phase. The illustrated device is called a Fresnel zone plate (FZP). An FZP focuses light to a distance $f$ by permitting and blocking the passage of light through regions that result in constructive and destructive interference, respectively, at the desired focal point.
2.3. Geometrical Phase

Another common approach for manipulating relative transmitted and reflected phases relies on the accumulation of a geometrical phase (GP), also called the Pancharatnam–Berry phase [106,107]. A phase of this type can be acquired between two light beams experiencing different variations in their polarization states while propagating over the same distance. Thus, the GP originates from the interplay between the polarization states, rather than being related to the dynamical phase acquired with propagation through a dielectric medium or the phase added to the scattered light by any scattering process. Because of the natural treatment of light polarization states that is available within the Poincaré sphere framework, this framework is commonly used to quantify GP accumulation. A typical design using the GP concept relies on laterally inducing different polarization variations of a propagating beam, as enabled by metasurface comprising nanoparticles with different orientations. As an example, we consider the case depicted in Fig. 4, in which light with right circular polarization (RCP) is shone on a surface consisting of nanorods with two different orientations. The first orientation is parallel to the x axis, and the second is rotated by an angle $\theta$. Thereafter, through the use of a polarizer and a quarter-wave plate, light with left circular polarization (LCP) is measured. By plotting on the sphere the paths of two beams (shown in red and green in Fig. 4) that have been subjected to different polarization changes, the relative GP can be determined through the following relation [93]:

$$\phi_g = \frac{1}{2} \Delta \Omega.$$  \hspace{1cm} (7)

**Figure 4**

Geometrical phase. Two paths with different polarization changes, resulting in the acquisition of a relative geometrical phase, are shown on the Poincaré sphere. Light in the RCP state is shone on a surface with two nanorod orientations, and the projection of the interference between the two beams in the LCP state is measured. The acquired relative phase between the two different paths is $\phi_g = \Delta \Omega/2$, where $\Delta \Omega$ is the solid angle encompassed by the paths.
where $\phi_g$ is the geometrical phase, and $\Delta \Omega$ is the solid angle encompassed by the two paths on the Poincaré sphere. To obtain some intuitive insight into the origin of the GP acquisition described above, we consider a single nanorod that is tilted at an angle $\theta$ with respect to the $x$ axis and illuminated by light with a general polarization state and qualitatively analyze how the GP shift between the final RCP and LCP states arises. For any polarization state, the incident light can be decomposed into LCP and RCP states. The greatest scattering amplitudes from the nanorod will be achieved for an electric field polarized parallel to the long axis of the nanorod. Thus, a nanorod rotated at angle $\theta$ will strongly react only when the incident light polarization is rotated by an angle $\theta$. If the RCP state has rotated by $\theta$ to be parallel to the long axis of the rod, then...
Examples of the linear shaping of light by using metasurfaces. (a) Artist’s view of a multilayered chromatically corrected metasurface lens. Through the vertical stacking of three different Fresnel zone plates based on plasmonic nanoparticles designed to operate in the red, green, and blue regimes of the visible spectrum, a spot of white light at the focal point can be generated under white light illumination. Images of the focal region of the lens are shown in (b) for three laser wavelengths of 450 nm, 550 nm, and 650 nm to show the chromatic aberration correction. (c) Finite-difference time-domain simulation of a wavefront created by light scattered from eight “V”-shaped nanoantennas with steady amplitude and phase variations over the full 2π range. Phase control is achieved via coupled resonances at the same nanoantenna. The resulting device can conceptually guide light in any desired direction through appropriate spacing of the scattering nanoantennas. (d) Scanning electron microscope image of a dielectric metasurface that operates as a blazed grating on the basis of the GP concept. (e) The measured diffraction patterns of the metasurface shown in (d) under illumination with RCP light (top), linearly polarized light (middle), and LCP light (bottom) at a wavelength of 550 nm. (f) Illustration of a generalized Brewster effect achieved with a silicon metasurface. By means of the interplay between electrical and magnetic dipoles, interference can be exploited to eliminate the reflection of s-polarized light while enhancing the reflection of p-polarized light, as shown in the simulated reflection curves. (g) Illustration of a dielectric metasurface designed for independent polarization and phase control at each unit cell. The metasurface is composed of elliptical amorphous silicon posts of the same height but different diameters and orientations. The orientation of each ellipse relative to the incident light polarization affords amplitude control, and the diameter of each ellipse determines the phase acquired during propagation through the surface. (a), (b) Reproduced from [108] under the terms of the Creative Commons Attribution 4.0 International License. (c) From Yu et al., Science 334, 333–337 (2011) [35]. Reprinted with permission from AAAS. (d),(e) From Lin et al., Science 345, 298–302 (2014) [70]. Reprinted with permission from AAAS. (f) From Kuznetsov et al., Science 354, aag2472 (2016) [91]. Reprinted with permission from AAAS. (g) Reprinted by permission from Macmillan Publishers Ltd.: Arbabi et al., Nat. Nanotechnol. 10, 937–943 (2015) [99]. Copyright 2015.

the LCP state must rotate by $2\pi - \theta$. Thus, the light components in the two circularly polarized states will differ in phase by $2\theta$. By varying the rod orientation angle $\theta$ between 0 to $\pi$, complete control over the full $2\pi$ range can be achieved while maintaining a steady transmission amplitude [70].

2.4. Examples of the Linear Shaping of Light by using Metasurfaces

Linear beam shaping by means of metasurfaces has been a fruitful area of research in the past few years and has been utilized for a variety of applications. Figure 5 shows several examples of recent works that have exploited the ability to control the phase and amplitude of scattered waves by means of either the response of a single nanostructure [Figs. 5(c), 5(f), and 5(g)] or the arrangement of an entire metasurface [Figs. 5(a), 5(b), 5(d), and 5(e)], for both plasmonic [Figs. 5(a)–5(c)] and dielectric [Figs. 5(d)–5(g)] metasurfaces. Phase control based on single-particle response has been achieved in these works either by operating in a nonresonant regime [Fig. 5(g)] or by exploiting the total impact of multiple resonances in the same nanostructure [Figs. 5(c) and 5(f)]. At the level of the arrangement of the entire metasurface, phase control has been achieved in these works by means of either a detour phase [Figs. 5(a) and 5(b)] or geometrical phase acquisition [Fig. 5(d)].
3. NONLINEAR OPTICAL INTERACTIONS WITH METASURFACES

The strong confinement of light to subwavelength regimes that occurs naturally in most metasurfaces also promotes enhanced nonlinear optical processes. Similarly to the response of linear metasurfaces, the amplitude, phase, and polarization of the local nonlinear response of the metasurfaces can be tuned by modifying the geometry. Therefore, in addition to the linear manipulation of light, manipulation of the effective nonlinear tensor can also be performed to tailor a functional nonlinear metasurface.

3.1. Local Nonlinear Polarization

Below, we will briefly elaborate on optical responses in nonlinear media. We will then focus on the fundamental aspects of optical nonlinearity that are essential for describing the nonlinear processes in metasurfaces. Further description can be found in [31].

The relation between the induced polarization of a material and the applied electric field is, in general, nonlinear. Their dependence, using the SI unit system, is usually described in the form of a Taylor expansion:

\[
P = \varepsilon_0 \chi^{(1)} E + \chi^{(2)} E^2 + \chi^{(3)} E^3 + \ldots \tag{8}
\]

where \( \chi^{(1)} \), \( \chi^{(2)} \), and \( \chi^{(3)} \) are the linear, quadratic, and cubic susceptibilities, respectively, and higher-order susceptibilities are not shown. The nonlinear terms in this expansion allow for the description of the generation of new frequencies and additional nonlinear phenomena that emerge from photon–photon interactions.

For example, for a general fundamental field composed of two different frequencies, \( \omega_1 \) and \( \omega_2 \), the quadratic nonlinearity induces oscillating polarizations at new frequencies, \( 2\omega_1, 2\omega_2, \omega_1 + \omega_2, \) and \( \omega_1 - \omega_2 \), as well as a DC polarization. In these cases, the nonlinear oscillating polarization takes the form of a second-order term, \( P^{(2)} = \chi^{(2)} E^2 \), which acts as a source for the wave equation that propagates the generated radiation in space and time:

\[
\nabla^2 E - \frac{n^2}{c^2} \frac{\partial^2 E}{\partial t^2} = \frac{1}{\varepsilon_0 c^2} \frac{\partial P^{(2)}}{\partial t}. \tag{9}
\]

The quadratic nonlinear processes are called three-wave mixing (TWM), in which frequency doubling, summation, and differentiation are termed second-harmonic generation (SHG), sum frequency generation (SFG), and difference frequency generation, respectively. The ability to manipulate a spatially nonlinear polarization by using a metasurface enables the control and shaping of the generated radiation. This is the underlying mechanism for the nonlinear shaping of light by using metasurfaces, which can be achieved either by controlling the shape of the incident field or by designing the local quadratic nonlinearity on the metasurface.

For example, for the case of SFG, the quadratic nonlinear interaction on a metasurface can be described in the frequency domain as follows:

\[
P_i^{(2)}(\omega_1 + \omega_2, x, y) = \varepsilon_0 \chi_{ijk}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2, x, y) E_j(\omega_1, x, y) E_k(\omega_2, x, y), \tag{10}
\]

where \( x \) and \( y \) are spatial coordinates along the metasurface and the \( z \) coordinate is absent, owing to the two-dimensional nature of metasurfaces. The nonlinear susceptibility \( \chi_{ijk}^{(2)} \) is a third-rank tensor that is usually defined by the material properties, and the participating frequencies. Here we define \( \omega_1 \) and \( \omega_2 \) as the fundamental frequencies and \( \omega_1 + \omega_2 \) marks the generated frequency [31].
It can be seen from Eq. (8) that in centrosymmetric materials, in which symmetry conditions locally require \( P(E) = -P(-E) \), all of the even-order terms in the expansion must disappear. Therefore, quadratic interactions, for example, cannot be observed in bulk centrosymmetric materials.

### 3.2. Sources of Nonlinearity in Metasurfaces

The great success of the development of linear metamaterials with unique properties \([35,68–75,91,99,108]\) led to the natural progression to exploration of nonlinear optical interactions with these new artificial compounds. In the following section, we will present a brief overview of the different sources of nonlinearities in metasurfaces. Most of the work done in the field of nonlinear metasurface, and specifically nonlinear shaping of light, is based on plasmonic metasurfaces. Therefore, in this tutorial we mainly concentrate on this type of metasurface. However, the field of dielectric nonlinear metasurfaces, which is now beginning to grow, holds within a great promise. For example, nonlinear metasurfaces composed of silicon nanodisk-based metamolecules \([110]\), Fano-resonant silicon structures \([111]\), and GaAs nanopillar resonators \([112]\) can be used to host the nonlinear interaction with lower loss and a flexible control of the bandwidth. As this field will continue to evolve, the majority of the concepts described in the following sections will be relevant for shaping of light also with dielectric nonlinear metasurfaces.

#### 3.2a. Integrating Metasurfaces with Conventional Nonlinear Materials

The most straightforward way to generate nonlinearities and to enhance the nonlinearity of a metasurface is to use conventional nonlinear materials to construct the metasurface or to embed the metasurface in a nonlinear platform. For example, implementing plasmonic nanoantennas inside a nonlinear crystal can enhance the overall effect, because the nonlinear material serves as a coupler between the linear and nonlinear modes. This phenomenon has been demonstrated, for example, in the forms of hole arrays and SRR arrays over GaAs as well as bowtie antennas over \( \text{LiNbO}_3 \) and more \([113–116]\). In addition to conventional nonlinear crystals, in which the nonlinear interaction is usually between virtual electronic states, it is possible to use resonant nonlinear materials to enhance the interaction \([31]\). This application has been shown, for example, by using multiquantum wells (MQWs) band-engineered to enhance specific TWM processes \([117–120]\). One of the factors limiting the use of strong nonlinearities in MQWs is that the transitions are polarized in the direction perpendicular to the layers, and therefore, they are not accessible for plane waves at normal incidence. To address this shortcoming, metasurfaces have been used to align the fundamental field polarization in the required direction, along with matching the plasmonic resonances and intersubband transitions, thus resulting in giant optical nonlinearities \([53,121–123]\) and nonlinear functionality, as depicted in Fig. 6(a).

#### 3.2b. Creating Artificial Quadratic Nonlinearity

Another important source of strong quadratic nonlinearity in metasurfaces is symmetry breaking at the interfaces of the nanostructures. It is well known that interfaces support strong quadratic nonlinearities \([128–130]\). In nanostructured materials, in which interfaces are abundant, there are very strong local quadratic nonlinearities, which can be further enhanced by field confinement. This enhancement is one of the motivations for using metasurfaces as artificial nonlinear materials. However, owing to the complex optical and plasmonic mode structure, it is important to understand the participating optical modes, their field overlap, and their radiation properties to enable the design of efficient nonlinear metasurfaces. Below, we present a model that captures the elementary physics of quadratic nonlinearity in metal nanostructures and
hence can be used to enhance and optimize the nonlinear interaction in these artificial optical elements.

In the context of plasmonic metamaterials, metals such as gold and silver are commonly used to construct the fundamental elements of these materials. These metals do not possess intrinsic symmetry breaking; consequently, their bulk quadratic susceptibility, $\chi^{(2)} = 0$. However, it has been specifically shown that quadratic processes involving field components that are normal to the surface are substantially stronger [131–135]. One common way to model the nonlinearity on a metal surface is by using

**Figure 6**

Different types of nonlinear metasurfaces. (a) Multiquantum well structure coupled to a plasmonic structure, resulting in a giant enhancement of the nonlinear optical interaction. (b) L-shaped metallic resonators. The arrangement of the orientations defines the nonlinear polarization interaction. (c) G-shaped resonators exhibiting optical chiral dichroism. (d) Triangular prism nanocavities. (e) Structure consisting of coupled resonators. The nonlinear field that is generated at the rod penetrates the discs and is emitted as light. The asymmetry of the structure controls the emission efficiency. (a) Reprinted with permission from [53]. Copyright 2016 Optical Society of America. (b) Reprinted with permission from Husu et al., Nano Lett. 12, 673–677 (2012) [124]. Copyright 2012 American Chemical Society. (c) Reprinted with permission from Valev et al., Nano Lett. 9, 3945–3948 (2009) [125]. Copyright 2009 American Chemical Society. (d) Reprinted with permission from Salomon et al., J. Phys. Chem. C 117, 22377–22382 (2013) [126]. Copyright 2013 American Chemical Society. (e) Reprinted with permission from Gennaro et al., Nano Lett. 16, 5278–5285 (2016) [127]. Copyright 2016 American Chemical Society.
a hydrodynamic model [136,137], in which the conduction electrons in the metal are regarded as forming an electron gas. In this case, the modified Euler’s flow equation takes the following form:

$$m_e n \left[ \frac{\partial v}{\partial t} + (v \cdot \nabla) v + \gamma v \right] = en(E + v \times H) - \nabla p,$$

(11)

where $v$ is the electron velocity; $E$ and $H$ are the electric and magnetic EM field components; $\gamma$, $n$, $e$, and $m_e$ are the scattering rate, density, charge, and effective mass, respectively, of the conduction electrons; and $p$ is the electron quantum pressure due to Pauli’s exclusion rule. The electron acceleration is represented by the first term on the left-hand side, the second term represents the convection of the electrons, and the third term represents the damping force. On the right-hand side, the first term represents the driving EM forces, and the second term represents the quantum pressure forces exerted on the electrons.

The polarization is derived from the electron motion as $\dot{P} = env$, along with the continuity relation $n = n_0 - \frac{1}{e} \nabla \cdot P$, where $n_0$ is the rest electron density. Thus, the flow equation takes the following form:

$$\ddot{P} + \frac{1}{n_0 e} [(\nabla \cdot \dot{P}) \dot{P} + (\dot{P} \cdot \nabla) \dot{P}] + \gamma \dot{P}$$

$$= \frac{n_0 e^2}{m_e} E - \frac{e}{m_e} E (\nabla \cdot P) + \frac{e}{m_e} \dot{P} \times H + \frac{k_B T (\nabla \cdot P)}{m_e}.$$

(12)

This equation can be solved in a perturbative manner, i.e., by spanning the different fields, $E$, $H$ and $P$, with oscillatory harmonic components. Solving for the fundamental polarization yields the dependence of the linear polarization on the driving fields. The same equation, solved for the nonlinear polarization, yields the dependence on the calculated linear polarizations, owing to the nonlinear terms in the relation. The nonlinear polarization acts as a source term in the wave equation, Eq. (9), for the radiation of light at new frequencies. The entire solution process is usually performed numerically [132,137–140], and the induced linear polarization gives rise to new generated frequencies through the surface currents. In the case of SHG, the surface current can be approximated to take the following form [137]:

$$K_{NL} = \frac{i \omega}{n_0 e} \left[ \hat{t} (P_\perp P_\parallel) + \hat{n} \frac{1}{2} \frac{3 \omega^2 + i \gamma}{2 \omega + i \gamma} (P_\perp)^2 \right].$$

(13)

where $\omega$ is the fundamental frequency; $\hat{n}$ and $\hat{t}$ are the unit vectors in the directions normal and transverse to the surface, respectively; and $P_\perp$ and $P_\parallel$ are the polarization components of the fundamental frequency (FF) field in the directions corresponding to $\hat{n}$ and $\hat{t}$, respectively.

The existence of strong surface nonlinearity does not guarantee that the nonlinear sources will radiate to the far field. In the case of macroscopic, yet subdiffraction-limit centrosymmetric structures, for each polarization element, there is an opposite polarization element that interferes destructively in the far field; consequently, the source does not radiate [135]. Therefore, breaking of this symmetry via the geometry of the nonlinear nanoparticles is also needed. This has been demonstrated both theoretically and experimentally for various structures, such as SRRs [44,45,137,141], L-shaped [90,124,142,143], T-shaped, chiral G-shaped [125,144], and triangular [145] nanoparticles, nonsymmetric dimers [146] and metamolecules.
and some of their complementary nanocavities [126,148,149]. A few examples are presented in Fig. 6. Figures 7(a) and 7(b) present the SH fields emitted from a bar antenna and an SRR antenna. In the case of the rod, two opposite dipoles are excited, which destructively interfere in the far field, whereas in the case of the noncentrosymmetric SRR, two dipoles are excited in the same direction and emit in a constructive manner to the far field. Cubic nonlinearity, however, does not require symmetry breaking at the level of either the crystal structure or the inclusion geometry. In fact, in addition to surface effects, cubic nonlinearity may occur simply because of conventional bulk $\chi^{(3)}$ properties. Consequently, cubic nonlinear effects have been

![Figure 7](image)

Models for the estimation of SHG efficiency. (a) Numerical simulation of the SHG from a nanorod in accordance with the hydrodynamic model of an electron gas. The SH mode acts as two opposite dipoles, which interfere destructively with each other and consequently do not radiate to the far field. (b) The SHG from an SRR simulated in the same way as in (a). Owing to the symmetry breaking, two SH dipoles along the arms are excited, radiate constructively, and emit to the far field. (c) Nonlinear scattering model for predicting the efficiency of SHG. The overlap integral of the participating radiating modes, i.e., $E(\omega)$ and $E(2\omega)$, is proportional to the effective nonlinear coefficient. (c) Reprinted by permission from Macmillan Publishers Ltd.: O’Brien et al., Nat. Mater. 14, 379–383 (2015) [141]. Copyright 2015.
observed from metasurfaces with simpler geometries, such as nanodisks, nanorods, surfaces with rectangular holes, bowtie antennas, and macroscopic gratings [81,110,150–159].

3.3. Harmonic Mode Matching

Nonlinear plasmonic nanoresonators interact strongly when excited at resonance. In this case, the field is confined to the nanoresonator, and the local nonlinearity is enhanced. This phenomenon has been used to introduce plasmonic-based nonlinear enhancement in a conventional nonlinear crystal embedded with a plasmonic metamaterial [160] and to enhance the cubic nonlinearity in nanocavities [152], and it is generally applicable to all nonlinear plasmonic metamaterials. Resonance occurs for specific spatial modes of a nanoresonator, and to achieve efficient nonlinear mode conversion, it is necessary to ensure good field overlap between the generated nonlinear local field and the radiating mode. This concept is described by the nonlinear scattering theory [141,161,162]. In SHG on a metal nanoresonator, for example,

\[ \chi^{(2)}_{\text{eff}} \propto \int_S \chi^{(2)}_{111} E^2_{1,\omega}(r) E_{1,2\omega}(r) dS, \]  

in which the effective quadratic nonlinear susceptibility \( \chi^{(2)}_{\text{eff}} \) is proportional to the mode mixing integrated over the surface of the nanoresonator. Here, \( \chi^{(2)}_{111} \) denotes the nonlinear tensor element that mixes the normal field components, \( E_{1,\omega} \) and \( E_{1,2\omega} \), which correspond to the fundamental and SH mode fields, respectively. Figure 7(c), taken from O'brien et al. [141], illustrates the concept of mode matching on SRRs for the enhancement of nonlinear emission.

When two or more plasmonic nanoparticles are in proximity, it is possible to harness the mode coupling and overlap in this multiresonant structure for nonlinear interaction enhancement, as has been demonstrated both theoretically and experimentally for a three-nanorods configuration [163], a V-shape and rod configuration [164], T-shaped dimers [146], and a nonsymmetric configuration consisting of a rod and two disks [127].

4. CONTROL OVER LOCAL NONLINEAR TENSOR ELEMENTS FOR THE NONLINEAR SHAPING OF LIGHT

The ability to tailor the wavefront of the light emitted from a nonlinear material is mostly based on the ability to change the spatial characteristics of the nonlinear interaction, i.e., the ability to locally modify the nonlinear susceptibility tensor. The matrix elements of this tensor are, in general, complex terms that set the amplitude and phase of each polarization component of the nonlinear field for any given fundamental field. In a conventional nonlinear material, the nonlinear tensor is determined by the atomic structure of the material, and its manipulation is very difficult. For example, for quadratic interactions, the most common method of locally controlling the \( \chi^{(2)} \) tensor is through the electric field poling of nonlinear ferroelectric materials. This method can locally invert the atomic structure and consequently effectively changes only the sign of the nonlinear coefficient. This method has been extensively used for controlling TWM interactions and as the basis for the development of nonlinear photonic crystals [26,27,32,165–169].

Nonlinear metasurfaces have several major advantages with regard to control over the nonlinear tensor. As explained above, their responses to an EM field are defined by their geometry, size and material, thus providing access to a broad and continuous spectrum of desired linear and nonlinear interactions. Furthermore, the same technique used to fabricate a uniform structure can be applied for the fabrication of any arbitrary combination of different types of inclusions in any kind of arrangement.
This technique allows for full spatial control over the nonlinear tensors of metasurfaces, in contrast to the case for conventional natural nonlinear crystals. Finally, three-dimensional nonlinear crystals have fundamental limitations regarding phase matching owing to the dispersion of the EM waves propagating through them. Because nonlinear metasurfaces effectively do not have a longitudinal length in the propagation direction, they are not subject to this disadvantage.

4.1. Control Over Amplitude

There are several ways to achieve local amplitude control of the effective macroscopic nonlinear coefficients. The most straightforward method is to control the local density of the meta-atoms. In the simplest approximation, the amplitude of the local average nonlinear coefficient will be proportional to the number of inclusions per unit area, $\chi_{\text{eff}}^{(2)} = n\beta$, where $n$ is the density of the meta-atoms and $\beta$ is their quadratic nonlinear polarizability. However, this simplified relation does not hold in certain regions, owing to near-field coupling between the LSP modes of neighboring meta-atoms [45,170], as presented in Fig. 8(a), and long-range coupling due to lattice effects. This coupling can lead to enhancement or degradation of the effective nonlinear coefficient [171–173].

Another method of controlling the amplitude through the interactions between meta-atoms is based on far-field interference to decrease the intensity of the emitted light.

**Figure 8**
When two nonlinear meta-atoms are separated by a sub-diffraction-limit distance, they act as a single emitter, emitting the sum of the two interfering fields. If both emitters are assumed to be identical, their emission would interfere constructively. However, shifting the relative phase between the emitters (as will be described in the following section) would result in intermediate interference, thus decreasing the emission amplitude proportional to $\cos(\Delta \phi / 2)$, where $\Delta \phi$ is the phase shift between the emitters [174]. It is important to note that the phase shift should be symmetric—i.e., the first meta-atom should be shifted by $\Delta \phi / 2$, and the second, by $-\Delta \phi / 2$—for the emission to be in phase with the rest of the metasurface.

At the level of a single meta-atom, in the simplified approach, the local nonlinear polarization depends quadratically on the field enhancement provided by the LSP modes in the inclusions, $P(2) \propto E_{local}^2 = (EF)^2 \times E_{inc}^2$, where $EF$ is the field enhancement factor associated with the LSP phenomenon and $E_{inc}$ is the incident field. By changing the nanoparticle morphology, the LSPR can be shaped to provide different local field enhancements. However, in this case, one must also consider the phase difference related to off-resonance excitations. In addition, as discussed earlier, the far-field nonlinear emission depends on the overlap between the fundamental and nonlinear modes of the meta-atom [Eq. (14)]. Therefore, through geometrical tuning of the structure, e.g., by modifying the ratio between the arms and the total length of an SRR [50,87,141,162] or the distance between dimers [155,157,163,175], one can vary the nonlinear efficiency of nanoparticles. This method of controlling the amplitude of the nonlinear coefficient is presented in Fig. 8(b).

4.2. Control Over Phase

Controlling the relative phase of emitted light is a powerful tool for the design of optical wavefronts. It is widely used in linear holography [19,176] and in conventional nonlinear beam shaping [27–30]. In the latter case, the electric field poling of...
ferroelectric materials is most commonly used to achieve binary control over the relative phase between poled and unpoled domains. However, the applicability of this technique is limited to only a few specific ferroelectric crystals whose atomic structures can be manipulated. The binary method of phase control has also been demonstrated in quadratic nonlinear metasurfaces by using the inversion asymmetry of the inclusions. Instead of inverting the orientation of the atomic structure, as is done in conventional nonlinear crystals, the orientation of the meta-atomic structure is modified. This inversion operation in SRRs, for example, has been shown to result in an opposite relative phase, as depicted in Fig. 8(c) [48,49].

In addition to the binary phase approach, continuous control over the nonlinear phase has been demonstrated by means of two main approaches. The first considers the phase accumulated through the linear interaction of each of the participating waves. As described in the previous section, in the case of single nanoresonators, a phase response of $0 - \pi$ around the resonance can be accumulated, depending on the operation frequency relative to the resonance frequency. In the case of a nonlinear polarization, the phase of the nonlinear output is described as follows [51]:

$$\phi^{(n)} = \sum_j \phi_j,$$

(15)

where $\phi_j$ is the phase imparted by the resonator to the jth component and $\phi^{(n)}$ is the total phase of the nth component of the nonlinear polarization acquired by the interacting fields. Thus, the phase space for light manipulation can be expanded. For example, for simple, single-resonance nanoresonators in the case of quadratic interaction, the full $2\pi$ range of manipulation is obtainable. This approach has been demonstrated for the case of four-wave mixing (FWM) in rectangular nanoholes with varying aspect ratios, in which $\omega_{FWM} = 2\omega_1 - \omega_2$, and the corresponding nonlinear phase was $\Delta \phi_{FWM} \approx 2\phi_1 - \phi_2$, as shown in Figs. 8(d) and 8(e) [51,152].

The phase gained by a nonlinear meta-atom can also be calculated in a numerical simulation. Thus, the construction of nonlinear phase and amplitude diagrams according to geometrical parameters of the meta-atom allow to plan a route in the geometrical parameter space which conserves the amplitude, but varies the phase, or vice versa [51,177]. In the case of quadratic nonlinearity, this method can be combined with inversion of the meta-atoms to gain additional phase of $\pi$ in order to get full control over the phase.

The second approach is based on a GP. As described in the previous section, light that is scattered by a resonator excited by circularly polarized light would gain a phase relative to a similar rotated resonator (see Subsection 2.3). In the nonlinear case, it has been shown that when an SRR, for example, is excited with circularly polarized FF light, it emits circularly polarized SH light of both handednesses. The light that is polarized with the same handedness as the excitation light would gain a phase equal to the rotation angle, whereas the SH light that is polarized with the opposite handedness would gain a phase of 3 times the rotation angle, as also presented in Fig. 8(f) [53,178,179]:

$$\chi_{RRR}^{(2)}(\Delta \varphi) = \chi_{RRR}^{(2)} e^{i\Delta \varphi}, \quad \chi_{LRR}^{(2)}(\Delta \varphi) = \chi_{LRR}^{(2)} e^{3i\Delta \varphi},$$

$$\chi_{RLL}^{(2)}(\Delta \varphi) = \chi_{RLL}^{(2)} e^{-3i\Delta \varphi}, \quad \chi_{LLL}^{(2)}(\Delta \varphi) = \chi_{LLL}^{(2)} e^{-i\Delta \varphi},$$

(16)

where R and L are the right and left circular polarization components and $\Delta \varphi$ is the rotation angle of the meta-atom. Similarly, it has been shown that the geometrical
phase also can be used to manipulate the phase of FWM phenomena. These mechanisms can be used to spatially tailor the nonlinear phase responses of metasurfaces.

4.3. Control Over Polarization

Polarization provides an important means of control over nonlinear interaction. Meta-atoms of different shapes enable the excitation of different polarization-dependent or polarization-independent modes [180]. As a consequence, the shapes and orientations of meta-atoms can be used to locally shape the full nonlinear tensor. For example, it has been shown that in an SRR, for an incident wave that is polarized parallel to the base of the SRR, the most efficiently generated nonlinear SH mode is cross polarized along the arms, owing to the strong linear interaction of the FF light with an LSP mode of the SRR when the FF light is polarized along the base. This interaction results in the creation of nonlinear currents along the arms of the SRR [see Fig. 7(b)]. The nonlinear currents radiate an efficient SH field that is polarized along the arms, thereby defining a strong $\chi^{(2)}_{yyx}$ component (in which the $x$ direction is parallel to the base and the $y$ direction is parallel to the arms) [44,137]. The most straightforward method of locally reshaping the nonlinear tensor is to rotate the meta-atom. In the case of an SRR, a rotation of $90^\circ$ transforms $\chi^{(2)}$ such that the prominent element is $\chi^{(2)}_{xyx}$. A more detailed structure of the nonlinear tensor can be achieved by using a more complex metamolecule. For instance, for a metamolecule that contains four L-shaped nanoantennas, each with a certain orientation, the polarization dependence of the metamolecule can be specifically tailored [see Fig. 6(d)] [124].

The case of circular polarization allows the use of rotational symmetry properties. As mentioned before, through the rotation of the meta-atom, a relative GP difference can be acquired. Moreover, if the meta-atom itself possesses a rotational symmetry, it is possible to control the type of polarization emitted with regard to the input polarization and the type of nonlinear interaction. For example, nanoantennas with C2 rotational symmetry allow third-harmonic generation (THG) with both left and right circular polarization, whereas nanoantennas with C4 symmetry do not allow THG with the same circular polarization. Through a similar mechanism, C3 symmetry causes the extinction of SHG with the same circular polarization as that used for excitation [145,158,178,181]. An additional characteristic of rotational symmetry is chirality. When the mirror image of a nanostructure does not coincide with itself, it is considered to be a chiral structure. One of the prominent properties of chiral structures is that they respond differently to right and left circular polarizations. In the context of nonlinear metasurfaces, chiral meta-atoms such as twisted-cross nanodimers [182], metamolecules with threefold symmetry constructed of triangular gold nanoprisms [147] [as seen in Fig. 8(g)], G-shaped and star-shaped nanoantennas [as shown in Fig. 6(c)] [183], and twisted arc structures [184] produce nonlinear light emission with higher intensity if the fundamental light is polarized with a handedness matching that of the chiral structure.

5. NONLINEAR DIFFRACTION—EXPLOITING THE COLLECTIVE NONLINEAR RESPONSE

5.1. Collective Effects on Nonlinear Metasurfaces

The assembly of base units into ordered arrays leads to nonlocal collective effects. It has been shown that the nonlinear conversion efficiency of such an array is dependent on the lattice constant of the array due to near-field and extended near-field interactions between the individual elements and dilution effects [170,185]. In the other regime, as mentioned above, when the distances between the nanoantennas are of the order of the wavelength, the condition for lattice resonances can be fulfilled. This collective phenomenon strengthens the excitation of the fundamental LSP mode when
tuned to the corresponding frequency, consequently enhancing the overall nonlinear efficiency [173]. When the lattice resonance is tuned to the emitted mode, the collective nonlinear interaction also results in a significant enhancement of nonlinear emission [172]. Recently, it has been shown that the effective nonlinear quadratic polarizability of the meta-atoms in such an array for the case of TWM and normal incidence exhibits behavior of the following form [172]:

\[
\beta_{\text{eff}}(\omega_3; \omega_1, \omega_2) = \frac{\beta_0(\omega_3; \omega_1, \omega_2)}{(1 - S(\omega_1)\alpha_s(\omega_1))(1 - S(\omega_2)\alpha_s(\omega_2)) - S(\omega_3)\alpha_s(\omega_3)),}
\]  

where \(\omega_1\) and \(\omega_2\) are the frequencies of the impinging waves; \(\omega_3 = \omega_1 \pm \omega_2\) is the generated nonlinear frequency; \(\alpha_s\) and \(\beta_0\) are the single-particle linear and quadratic polarizabilities, respectively; and \(S(\omega)\) is the retarded dipole sum, which encompasses the collective array response. This expression for the nonlinear polarizability \(\beta_{\text{eff}}\) takes a form similar to that of the quadratic susceptibility of the local Miller’s rule [186], whereas the localized resonances that give rise to the enhanced \(\chi^{(2)}\) in Miller’s rule are replaced with the linear and nonlinear surface lattice resonance conditions, \(\text{Re}\{S(\omega_k)\alpha_s(\omega_k)\} = 1\) (for \(k = 1, 2\)) and \(\text{Re}\{S(\omega_k)\alpha_s(\omega_2)\} = 1\), respectively.

Equation (17) is derived by extending the coupled dipole approximation to the nonlinear regime for the case of TWM. When either the separation between the meta-atoms or the modulation period of the amplitude or phase of the nonlinear emission is larger than the output wavelength, diffraction effects start to govern the far-field emission patterns, and such metasurfaces can be analyzed as nonlinear photonic crystals.

5.2. Nonlinear Photonic Crystals

The theory of nonlinear photonic crystals for quadratic nonlinear materials was introduced in 1998 by Berger [165]. It has been shown that for a nonlinear photonic crystal in which \(\chi^{(1)}\) is uniform and \(\chi^{(2)}\) is periodically modulated, the momentum in the nonlinear interaction must be conserved up to a reciprocal lattice translation. For a three-dimensional nonlinear photonic crystal, the nonlinear momentum conservation relation becomes

\[
k_G = \sum_i k_i + G,
\]

where the \(k_i\) are the wave vectors of the generating waves, \(k_G\) is the wave vector of the generated wave that can be matched in the nonlinear interaction on the lattice, and \(G = m_1K_1 + m_2K_2 + m_3K_3\) is the lattice momentum vector, where the \(K_{1,2,3}\) are the primitive reciprocal lattice vectors and the \(m_{1,2,3}\) are integers [165]. When a nonlinear interaction obeys momentum conservation, in addition to its inherent energy conservation, the interaction becomes efficient. Owing to the dispersion in natural materials, which leads to inherent momentum mismatch in nonlinear interactions, nonlinear photonic crystals are extensively used to achieve efficient frequency conversion [32,165,166].

The theory of nonlinear photonic crystals can be reduced to two-dimensional optical structures, i.e., for nonlinear metasurface-based photonic crystals. In these cases, the periodic modulation can be applied by subwavelength changes in the meta-atomic structure (e.g., phase control by rotation of the nonlinear meta-atom) or by subwavelength changes in the inter-meta-atoms distances. However, in order to apply the concepts of nonlinear photonic crystals for shaping the emitted light, generally, the modulation period must be longer than the emitted wavelength. Such modulation period defines the available reciprocal lattice vectors that allow to fulfill the
momentum matching condition for the generated light, as will be demonstrated in the following section. The two-dimensional nature of metasurfaces allows periodic modulations only in the transverse directions. Along with the assumption of incident fields in the form of plane waves, i.e., \( E_ω(\mathbf{r}) = E_ω e^{i \mathbf{k} \cdot \mathbf{r}} \), the expression for the nonlinear polarization generated in the SFG process presented in Eq. (10) can be written as

\[
P_{ω_3}(x, y, z_0) = ε_0 χ^{(2)}(x, y, z_0)E_{ω_1}E_{ω_2} e^{i(k_1 + k_2)·\mathbf{r}},
\]

where \( z_0 \) is the position of the metasurface. In the case of a one-dimensional periodic modulation of the quadratic nonlinear coefficient, for example, the coefficient can be expanded into a Fourier series of the following type:

\[
χ^{(2)}(x, y) = χ^{(2)}(x) \sum_m a_m e^{i \frac{2 \pi m}{\Lambda} x},
\]

where \( a_m \) is the \( m \)th Fourier component of the modulation function and \( \Lambda \) is the lattice constant. Thus, in this case, each term effectively adds a phase of \( \frac{2 \pi m}{\Lambda} \) to the \( m \)th component of the spatial nonlinear polarization. This addition is equivalent to the addition of the transverse momentum \( mK_x = \frac{2 \pi m}{\Lambda} \) by the lattice, thus leading to nonlinear diffraction. For the case of SFG, for example, the nonlinear polarization takes the following form:

\[
P_{ω_3}(x, y, z_0) = ε_0 χ^{(2)} \sum_m a_m E_{ω_1}E_{ω_2} e^{i(k_1 + k_2 + mK_x)·\mathbf{r}},
\]

where, owing to the two-dimensionality of the metasurface, the \( z \) component adds only a uniform phase to the interaction. Therefore, only the parallel components of the momentum are included in the momentum conservation, similarly to the nonlinear Raman–Nath relation [187]:

\[
k_{3,m}^∥ = k_1^∥ + k_2^∥ + mK_x
\]

and

\[
k_{3,m}^⊥ = \sqrt{\left(n(ω_3)\frac{ω_3}{c}\right)^2 - |k_{3,m}^∥|^2},
\]

where \( n(ω_3) \) is the refractive index of the medium of propagation. The result is a specific angle in which the nonlinear radiation will propagate. Because the metasurface has a deep subwavelength thickness, in principle, the nonlinear emission is symmetric both forward and backward. Notably, in the case of different refractive indices in the different directions, the amplitude of emission and the corresponding wave vectors are affected. In the case of incident beams that are normal to the unmodulated crystal axis, the diffraction angle is given by

\[
\sin θ_{SFG,m} = \frac{k_{3,m}^∥}{|k_3|} = \frac{c}{n(ω_3)} \frac{k_1^∥ + k_2^∥ + mK_x}{ω_1 + ω_2}.
\]

For the case of SHG from a metasurface in a homogenous environment, this expression can be reduced to
\[
\sin \theta_{2,m} = \sin \theta_1 + m \frac{\lambda_1}{2\Lambda},
\]

as is also illustrated in Figs. 9(a)–9(c).

These concepts have been demonstrated for the case of nonlinear metasurfaces. Segal et al. [48] have reported the deployment of metasurfaces consisting of arrays of periodically inverted SRRs to obtain a nonlinear photonic crystal suitable for the SH radiation. The expected nonlinear diffraction pattern has been observed from the one-dimensionally periodic structure, as depicted in Figs. 9(d)–9(f). The concept of nonlinear metasurface photonic crystals (NLMPCs) for controlling the direction of emission has also been demonstrated for SHG and FWM by using binary modulation of the nonlinear coefficient [48,49,149] and phase gradient metasurfaces [51,53,177–179].

Because both the one-dimensional and two-dimensional patterns follow Raman–Nath diffraction [Eq. (25)], the excitation wavelength and the lattice constant can be varied.

Figure 9

Nonlinear diffraction. (a) Periodic modulation of the effective nonlinear coefficient can be achieved by a periodic variation in the orientation of the SRRs. As a result, the reciprocal space is spanned by the lattice momentum vectors. (b), (c) Two photons with momentum \(k_1\) at (b) normal incidence and (c) oblique incidence result in SHG emission that complies with momentum conservation of the sum of the momenta of the incident photons and the lattice momentum, in agreement with Raman–Nath diffraction. (d) SEM image of a nonlinear metasurface photonic crystal consisting of periodically inverted SRRs. (e) Fourier space imaging of SH emission from the sample depicted in (d), showing diffraction corresponding to the periodicity and wavelength. (f) All-optical scanning up to \(\theta_{\text{SH}} = 30^\circ\) obtained in accordance with Eq. (25) by adjusting the wavelength and period (i.e., \(K_{\lambda}\)). Reprinted by permission from Macmillan Publishers Ltd.: Segal et al., Nat. Photonics 9, 180–184 (2015) [48]. Copyright 2015.
to adjust the diffraction angle, i.e., the beam position in the Fourier plane, thereby enabling all-optical control of the deflected beams [Fig. 9(f)]. In addition to one-dimensional modulations, two-dimensional NLMPCs with square and hexagonal lattice structures have been shown to generate diffraction patterns with fourfold and sixfold symmetries, respectively. A quasi-periodic Penrose tiling structure has been shown to produce 10-fold symmetry in the deflected beams. These structures are depicted in Figs. 10(a)–10(d).

6. NONLINEAR LIGHT SHAPING INTO BEAMS AND ARBITRARY IMAGES

6.1. Nonlinear Focusing of Light

The concept of diffraction from a periodic structure can be extended to the creation of nonlinear metasurface lenses. The angle of diffraction from each point on the metasurface is defined by the local periodicity; therefore, it is possible to engineer the detour phases from all points on the metasurface to interfere constructively at a focal point. This type of lens, known as a Fresnel zone plate, has been used to concentrate the SH emission from a NLMPC, in its binary form [48]. The NLMPC is formed of rings of SRRs with alternating orientations. The radius of the $n$th ring obeyed

$$r_n = \sqrt{n f \lambda_{SH} + \frac{n^2 \lambda_{SH}^2}{4}},$$

where $\lambda_{SH}$ is the SH wavelength and $f$ is the focal length. The SHG Fresnel zone plate and its behavior are depicted in Figs. 11(a)–11(b).

Another method of concentrating nonlinear light emission is to imitate the phase front of a spherical lens. Using the phase control achieved by varying the aspect ratios of rectangular nanocavities, a nonlinear metasurface has been designed to generate FWM wavefronts with a phase obeying $\phi(r) = \sqrt{r^2 + f^2}$ [51]. The FWM light generated from this metasurface has been concentrated to a nearly diffraction-limited

![Figure 10](image)

Two-dimensional metasurface-based nonlinear photonic crystals. (a) SEM image of a square lattice nonlinear metasurface, with the corresponding square SH diffraction pattern shown in (b). (c) SEM image of a nonlinear hexagonal lattice and (d) its corresponding hexagonal SH diffraction pattern. Reprinted by permission from Macmillan Publishers Ltd.: Segal et al., Nat. Photonics 9, 180–184 (2015) [48]. Copyright 2015.
spot. The metasurface and its nonlinear emission intensity profile are depicted in Figs. 11(c)–11(d).

Nonlinear focusing with metasurfaces has also been extended to the dual-layer case, in which two focusing nonlinear metasurfaces operating at the THG frequency, each designed for a different orthogonal polarization and a different focal length, have been fabricated one on top of the other, creating a functional polarization-dependent nonlinear lens [52].

6.2. Nonlinear Image Encoding on a Metasurface

As discussed above, the possibilities of light shaping are much vaster than the simple diversion and focusing of the generated light. Harnessing the possibility to locally control the polarization, amplitude, and phase of a nonlinear point source on a sub-wavelength scale enables the formation of structures for complex nonlinear beam shaping and nonlinear holographic devices. In addition, regardless of the propagation of the light from such a metasurface, several methods have been presented for forming a nonlinear image on the metasurface, i.e., directly in the optical near field.

Figure 11

Nonlinear metasurfaces based lenses. (a) Nonlinear metasurface Fresnel zone plate composed of gold SRRs. (b) SH imaging of the metasurface depicted in (a) and the plane 1 mm away from the surface, showing the focusing of the SH light to a focal point. (c) SEM image of an FWM lens metasurface composed of rectangular nanoholes, imitating the phase front of a lens. (d) Intensity profile of FWM light emitted from the structure depicted in (c), showing the focusing of the light. (a), (b) Reprinted by permission from Macmillan Publishers Ltd.: Segal et al., Nat. Photonics 9, 180–184 (2015) [48]. Copyright 2015. (c), (d) Reproduced from [51] under the terms of the Creative Commons Attribution 4.0 International License. With copyright permission.
Watermarking using a chiral nonlinear metasurface has been demonstrated by paving the surface with shapes made of chiral nonlinear meta-atoms \[147,184\]. Under excitation with circularly polarized light, the radiated image takes the form of the imprinted shape, whereas excitation with the opposite polarization results in the complementary image, as shown in Fig. 12(a). A more elaborate scheme also allows for continuous amplitude control by defining each pixel as a combination of two meta-atoms with a geometrical phase difference due to a relative rotation angle, thus resulting in local interference that modifies the SH amplitude, as depicted in Fig. 12(b) \[174\]. This method allows to control the SH amplitude continuously, and encode SH images on the metasurface, as shown in Fig. 12(c). While a desired nonlinear

Figure 12

Nonlinear image encoding. (a) Nonlinear watermarking in an arbitrary shape using chiral metamolecules as pixels. Under excitation with right circularly polarized light, the desired arbitrary image is visible in the SH emission, whereas under excitation with left circularly polarized light, its complementary image is shown. (b) Nonlinear metamolecules consisted of two identical threefold rotational symmetry nonlinear meta-atoms. The relative rotation angle between the meta-atoms defines a different GP at the SH, and due to interference between the emitted waves, an effective SHG amplitude. (c) Illustration of a nonlinear metasurface encoded with metamolecules as described in (b) to give an arbitrary image on the metasurface with SH light. (a) Reprinted with permission from Kolkowski et al., ACS Photonics 2, 899–906 (2015) \[147\]. Copyright 2015 American Chemical Society. (b), (c) Reprinted with permission from Walter et al., Nano Lett. 17, 3171–3175 (2017) \[174\]. Copyright 2017 American Chemical Society.
light image is constructed as a result of the manipulation of the metamolecules along the metasurface, the linear images, i.e., the reflected and transmitted light, both in the FF and the SH frequency are not affected by this modulation [174].

6.3. Nonlinear CGH Metasurfaces for Far-Field Beam Shaping

CGH methods can also be applied for nonlinear beam shaping in general [30] and for nonlinear metasurfaces specifically. Generally, CGH nonlinear light shaping techniques rely on the ability to locally control the nonlinear phase and amplitude values in a continuous range. However, binary control of the nonlinear susceptibility can also be used for designing the wavefront of nonlinear diffracted light. In these cases, the binary computer-generated hologram technique, first presented by Lee [19], can be implemented. Through this method, the desired wavefront is encoded over a spatial carrier frequency. For example, in the quadratic case, to achieve a phase front \( \varphi(x, y) \) and an amplitude front \( A(x, y) \), the binary nonlinear susceptibility function is as follows [30]:

\[
\chi^{(2)}(x, y) = \chi^{(2)} \text{sign} \left\{ \cos \left[ \frac{2\pi x}{\Lambda} - \varphi(x, y) \right] - \cos [\pi q(x, y)] \right\}, \tag{26}
\]

where \( A(x, y) = \sin(\pi q(x, y)) \) and \( \Lambda \) is the period of the spatial carrier frequency. The spatial carrier frequency enables the separation of the beam into different diffraction orders by virtue of the different Fourier components of the square-wave function given by the binary modulation. Each diffraction order carries a phase front \( m \varphi(x, y) \), where \( m \) is the diffraction order index, as has been demonstrated for the shaping of the SHG emission from SRR-based nonlinear metasurfaces [50], in which nonlinear susceptibility modulation has been applied by inverting the SRR orientation to follow a general nonlinear tensor mapping of

**Figure 13**

Nonlinear binary CGH beam shaping. (a) Binary quadratic susceptibility map for the generation of an Airy beam hologram following Eq. (27). (b), (c) SEM image of a nonlinear metasurface for the generation of an Airy beam as in (a), consisting of SRRs. (d) Simulated and (e) measured SH Airy beams generated by the metasurface presented in (b), (c). Reprinted with permission from Keren-Zur et al., ACS Photonics 3, 117–123 (2016) [50]. Copyright 2016 American Chemical Society.
\[ \chi^{(2)}(r) = \chi_{SRR}^{(2)} \text{sign} \left\{ \cos \left[ \frac{2\pi x}{\Lambda} - \varphi(r) \right] \right\}. \] (27)

In that work, two examples have been presented: a vortex beam with an orbital angular momentum of \( \hbar l \) and a phase front of \( \varphi(r, \phi) = l\phi \), produced by a grating structure with a fork singularity, and an Airy beam with a phase front of \( \varphi(x, y) = f_c y^3 \), also depicted in Figs. 13(a)–13(c). The metasurfaces have been demonstrated to generate the desired beam shapes at the diffraction orders with a close match to simulations, as seen in Figs. 13(d) and 13(e). The beams were diffracted at angles consistent with the Raman–Nath condition [Eq. (25)]. Importantly, opposite diffraction orders carried beam profiles with opposite phases; therefore, the orbital angular momenta of the diffracted vortex beams were opposite, and the acceleration directions of the Airy beams were opposite as well. This method can be used to shape nonlinear emission in the far field to any type of beam shape or arbitrary image.

6.4. Perfect Nonlinear Beam Shaping

The direct nonlinear generation of pure eigenmodes of the paraxial Helmholtz equation can also be achieved with nonlinear metasurfaces. Doing so requires continuous control of the spatial distribution of the nonlinear phase and amplitude generated on the metasurface. This type of nonlinear beam shaping method is unique to metasurfaces, because the absolute value of the nonlinear susceptibility of a conventional nonlinear material is fixed.

This concept has been demonstrated in metasurfaces designed for the direct nonlinear generation of the first-order Hermite–Gauss mode from a nonlinear metasurface.

Figure 14

Perfect nonlinear beam shaping. (a) Slow variation of the geometrical structure of the meta-atoms along the metasurface, consistent with Fig. 8(b), in combination with inversion of their orientation, results in a linearly varying quadratic susceptibility. (b) SH Hermite–Gauss beam as imaged from the metasurface illustrated in (a) when excited with a Gaussian FF beam, and (c) the far-field image of the propagated beam, which maintains the same form. Reprinted with permission from Keren-Zur et al., ACS Photonics 3, 117–123 (2016) [50]. Copyright 2016 American Chemical Society.
consisting of SRRs. The metasurface has been designed to have a susceptibility tensor amplitude that linearly varied from \(\chi_{\text{max}}^{(2)}\) to \(\chi_{\text{max}}^{(2)}\), i.e., \(\chi^{(2)}(x,y) = \frac{\chi_{\text{max}}^{(2)} x}{a}\) for a surface width of \(a\) such that \(x \in [-\frac{a}{2}, \frac{a}{2}]\). The change in the amplitude of the susceptibility has been achieved through geometrical tuning of the dimensions of each of the SRRs along the \(x\) axis, and the sign change has been achieved by flipping the SRRs [Fig. 14(a)]. After the excitation of the nonlinear metasurface with a Gaussian beam of radius \(w\) at the surface, the following nonlinear polarization profile is generated:

\[
P^{(2)}(x,y) = \frac{\epsilon_0\chi_{\text{max}}^{(2)}}{a} E_0^2 \frac{x^2 + y^2}{w^2},
\]

which generated an exact nonlinear free-space Hermite–Gauss (0, 1) EM mode, as can be seen in both the optical near-field image and the optical far-field image shown in Figs. 14(b) and 14(c), respectively. This method can be extended to the generation of other beam types, such as higher-order Hermite–Gauss modes, simply by constructing the spatial dependence of the nonlinear tensor to follow a two-dimensional Hermite polynomial. Additionally, in a similar manner to linear vortex beam shaping, a nonlinear metasurface can be designed with a geometrical phase that exhibits a linear dependency on the angle relative to the center of the metasurface to generate an eigenmode of a vortex beam with the following wavefront:

\[
P^{(2)}(r, \phi) = \epsilon_0 \chi_{\text{max}}^{(2)} E_0^2 e^{-\frac{x^2 + y^2}{w^2} + id\phi}.
\]

### 6.5. Arbitrary Shaped Nonlinear Holograms

The concept of nonlinear holography can be extended to the general case of arbitrary shapes, as has been demonstrated with metasurfaces consisting of V-shaped gold nanoantennas for THG [52]. Each antenna was designed to emit at a certain phase and a constant intensity through the specific design of the length of the V-shaped arms and their relative angle. Polarization dependence was also considered and was controlled by tuning the angle of the nanoantenna relative to the metasurface axis. Mapping the metasurface with these antennas in accordance with a specific CGH has enabled the THG of arbitrary shapes in the far field. Moreover, by virtue of the polarization dependence and a careful analysis of the response of the nanoantennas to an unwanted polarization, a dual-polarization nonlinear hologram has been presented [also shown in Figs. 15(a) and 15(b)], showing two different holograms for two orthogonal polarizations. A multilayered metasurface of this type has similarly been used to generate three different nonlinear holograms, each generated by a different polarization (0°, 45°, 90°) and designed for a different focal distance, as presented in Fig. 15(c). These results demonstrate the possibility of creating polarization-controlled three-dimensional nonlinear holograms by extending the metasurface concept to multilayered designs.

Another approach for generating several holograms from a single nonlinear metasurface has been demonstrated by Ye et al. [54]. In that study, two different SHG holograms and a linear hologram were encoded on the basis of a geometrical phase. A metasurface consisting of SRRs was excited by circularly polarized light, and each of the nonlinear holograms was carried by a different circular polarization. The metasurface was also designed to interact with a linear beam to generate a different hologram at the fundamental frequency [54]. This concept is illustrated in Figs. 15(d)–15(e). Such a versatile design is possible because of the existence of multiple degrees of freedom. First, in most holographic imaging, only the intensity
profile of the image is desired, whereas the phase is of less importance. Second, an FF beam does not differ between a certain SRR and its inverted form, whereas the SH output carries a relative phase of $\pi$. Thus, local linear geometrical phase control is possible over a range of $2\phi$ for the FF beam (where $\phi$ is the SRR angle). Third, the nonlinear phase generated by a circular polarization of type $\sigma$ can be separated into a SHG component of circular polarization $\sigma$ with a phase of $3\phi$ and an additional SHG component of polarization $-\sigma$ that carries a phase of $3\phi$, as described in Eq. (16). As a result, the phase of the $+\sigma$ SHG component has a one-to-one relationship with the SRR angle $\phi$, whereas for the $-\sigma$ SHG component, the three orientation angles of $\phi$ and $\phi \pm \frac{2\pi}{3}$ yield the same nonlinear phase of $3\phi$. Incorporating these
emission phases for each separate beam into a CGH algorithm enables the construction of a wavelength-spin multiplexing holographic metasurface of this kind.

7. CONCLUSION AND OUTLOOK

The emerging field of nonlinear metasurfaces has been attracting a lot of attention in the recent years thanks to the vast range of possibilities it holds within and the many challenges that yet remain unsolved. For more information we advise the readers to turn to additional comprehensive reviews in the topic of nonlinear metasurfaces [46,47,57,188,189]. In this paper, we reviewed the recent progress in the linear and nonlinear generation and shaping of light by using metasurfaces. We specifically presented various mechanisms that enable control over the local linear and nonlinear interactions on such metasurfaces. We discussed the additional possibilities in the nonlinear case compared with the linear case, which stem from photon–photon interactions, and showed how to use them for the nonlinear shaping of light. Such nonlinear shaping of light can be useful for imaging and data communication, among other applications. The integration of these capabilities into engineered nanometer-scale layers and the fabrication of such structures over large surface areas by using standard techniques make these concepts interesting for the development of integrated nonlinear devices. We also described the advantages of using metasurfaces rather than conventional nonlinear materials for the control of single subwavelength pixels. Because each pixel can be engineered separately to emit with a certain amplitude, phase, radiation pattern, and polarization and with a specific spectral response, the toolbox for nonlinear light shaping by using metasurfaces is significantly broader than the shaping abilities afforded by conventional materials. The low conversion efficiencies achieved to date with metasurfaces remain a challenge for most practical applications; however, their high functionality can be already applied for various sensing applications. In addition, several works have demonstrated how to considerably improve the conversion efficiency by using, for example, metasurfaces integrated on quantum wells or multilayered devices. The concepts described here for the shaping of light by using nonlinear plasmonic metasurfaces can also be extended to nonlinear dielectric metasurfaces. In that case, the nonlinear interaction can be dominated by scattering with negligible absorption. Together with the evolution of 3D nanoscale fabrication techniques and developments in the field of active metamaterials, this
direction of study may lead to the realization of large-scale, actively tunable, nonlinear metamaterials that can be designed for specific applications.

**FUNDING**

Israel Science Foundation (ISF) (1331/13, 1433/15); H2020 European Research Council (ERC) (715362, 639402).

**ACKNOWLEDGMENT**

T. E. and H. S. acknowledge support by the ERC and by the ISF. S. K. acknowledges support from the Tel Aviv University Center for Renewable Energy President Scholarship for Outstanding Ph.D. Students and support from the Israel Ministry of Science, Technology and Space Scholarship for Ph.D. students in the applied sciences.

**REFERENCES**

76. C. F. Bohren and D. R. Huffman, Absorption and Scattering of Light by Small Particles (Wiley, 2004).


Shay Keren-Zur was born in Jerusalem, Israel. He received his B.Sc. degree (cum laude) in physics and chemistry from the Hebrew University in Jerusalem in 2011. He received his M.Sc. degree from the Department of Physics in the Hebrew University of Jerusalem in 2013, studying the conduction and photo-conduction of semi-conductors nanocrystals in the group of Prof. Oded Millo. He continued toward his Ph.D. degree in the Physical Electronics Department in the Electrical Engineering School at Tel Aviv University, Israel. In his Ph.D. research, in the group of Prof. Tal Ellenbogen, he studies quadratic nonlinear interactions in metasurfaces and methods to use them for light manipulation.
Lior Michaeli was born in Karmiel, Israel. He received his B.Sc. degree (cum laude) in physics from Tel-Aviv University in 2014. Since 2014, he has been studying toward his Ph.D. in the direct track in the Physics Department of Tel-Aviv University. In his Ph.D. research, under the supervision of Prof. Tal Ellenbogen and Prof. Haim Suchowski, he studies the collective linear and nonlinear coherent interactions in nanoparticle arrays.

Haim Suchowski is an assistant professor at the Department of Condensed Matter Physics, the School of Physics and Astronomy, Tel Aviv University. He performed his postdoctoral research at the University of California, Berkeley (2014), and his Ph.D. at the Weizmann Institute of Science (2011). He holds a B.A. in physics (2004) and a B.Sc. in electrical engineering (2004) from Tel Aviv University, and a M.Sc. in physics (2006) from the Weizmann Institute of Science. His research focuses in exploring ultrafast dynamics in condensed matter physics, plasmonic nanostructures, silicon photonics, and 2D materials. He also performs research in quantum coherent control of atoms and molecules with ultra-short laser pulses, and analogous schemes in nonlinear optics. He has 31 articles and six patents. He received the Fulbright postdoctoral fellowship and was awarded the ERC grant for his project “MIRAGE 20-15.”

Tal Ellenbogen is an associate professor at the Department of Physical Electronics in Tel-Aviv University. He obtained a B.Sc. degree in physics and computers, a M.Sc. degree in physics, and a Ph.D. degree in nonlinear optics from Tel-Aviv University. He then joined the School of Engineering and Applied Sciences at Harvard University as a postdoctoral fellow to work on nanoplasmonics and excitonics. During his postdoctoral fellowship he was also a member of The Center for Excitonics in MIT. He is now the head of the Laboratory for Nanoscale Electro Optics at Tel-Aviv University. His research interests are in the field of nanoscale light matter interaction.