A review of gap-surface plasmon metasurfaces: fundamentals and applications

Fei Ding*, Yuanqing Yang, Rucha A. Deshpande and Sergey I. Bozhevolnyi

Abstract: Plasmonic metasurfaces, which can be considered as the two-dimensional analog of metal-based metamaterials, have attracted progressively increasing attention in recent years because of the ease of fabrication and unprecedented control over the reflected or transmitted light while featuring relatively low losses even at optical wavelengths. Among all the different design approaches, gap-surface plasmon metasurfaces – a specific branch of plasmonic metasurfaces – which consist of a subwavelength thin dielectric spacer sandwiched between an optically thick metal film and arrays of metal subwavelength elements arranged in a strictly or quasi-periodic fashion, have gained awareness from researchers working at practically any frequency regime as its realization only requires a single lithographic step, yet with the possibility to fully control the amplitude, phase, and polarization of the reflected light. In this paper, we review the fundamentals, recent developments, and opportunities of gap-surface plasmon metasurfaces. Starting with introducing the concept of gap-surface plasmon metasurfaces, we present three typical gap-surface plasmon resonators, introduce generalized Snell’s law, and explain the concept of Pancharatnam-Berry phase. We then overview the main applications of gap-surface plasmon metasurfaces, including beam-steerers, flat lenses, holograms, absorbers, color printing, polarization control, surface wave couplers, and dynamically reconfigurable metasurfaces. The review is ended with a short summary and outlook on possible future developments.

Keywords: gap-surface plasmon metasurfaces; Pancharatnam-Berry phase; flat optical elements; polarization control; dynamically reconfigurable metasurfaces.

1 Introduction

The capability of manipulating light at will is tantalizingly attractive, promising numerous applications. Conventional methods to mold the flow of light are typically relying on gradually accumulated amplitude, phase, and polarization changes during light propagation [1], with the resulting devices featuring certain curved surfaces and complex shapes. These bulky configurations do not comply with current trends aiming at very dense integration and miniaturization in photonics/plasmonics.

In recent years, metasurfaces, i.e. thin planar arrays of resonant subwavelength elements arranged in a periodic or aperiodic manner, have attracted progressively increasing attention and become a rapidly growing field of research because of their unprecedented control over electromagnetic (EM) fields by modifying boundary conditions for impinging waves [2, 3]. Generally, metasurfaces significantly benefit from their simplified fabrication process compared with their three-dimensional (3D) analog – metamaterials that are usually made up of complicated 3D artificial nanostructures processing exceptional EM properties [4–6]. As such, metasurfaces are promising for on-chip integration in the area of photonics and plasmonics owing to their planar profiles. Among all the metasurfaces, gap-surface plasmon metasurfaces (GSPMs) – a specific branch of plasmonic metasurfaces – which consist of a subwavelength thin dielectric spacer sandwiched between an optically thick metal film and arrays of metal subwavelength elements arranged in a strictly or quasi-periodic fashion, have gained awareness from researchers working at practically any frequency regime as its realization only requires a single lithographic step, yet with the possibility to fully control the amplitude, phase, and polarization of the reflected light.

As metasurfaces comprise a rapidly growing field of research, several excellent review articles can be found,
some presenting a comprehensive review of the filed [3, 7–16], while others are mainly focused on a certain type of metasurfaces [17–21] or specific branches of applications [22–26]. In this review, we review the recent progress of GSPMs during the past few years and attempt to provide our prospective on this specific branch of metasurfaces. The rest of this paper is organized in the following sections. In Section 2, we briefly discuss and classify the typical gap-surface plasmon (GSP) resonators and then introduce the concept of generalized Snell’s law and Pancharatnam-Berry (PB) phase. Section 3 is devoted to the applications of GSPMs, including beam-steers, flat lenses, holograms, absorbers, color printing, polarization control, surface wave couplers, and dynamically reconfigurable metasurfaces. Finally, we summarize and provide perspective for future developments in Section 4.

2 Fundamentals

2.1 GSP resonator

Surface plasmon polaritons (SPPs) are EM excitations propagating along and tightly bound to metal-dielectric interfaces, which arise from the coupling of EM field in dielectric and collective electron oscillations in metal [27]. Owing to the unique properties of subwavelength confinement and large field enhancement, SPPs have motivated numerous potential applications, ranging from biochemical sensing [28], quantum optics [29], and information communication technologies [30] to sustained energy sources [31]. However, in the long-wavelength part of the visible spectrum, SPPs are quite close to the light line and unwieldy to be utilized to achieve strongly enhanced local field. To realize strong field enhancement at any wavelength, one effective method is to place two metal-dielectric interfaces close to each other, which introduces the near-field coupling (NFC) between the SPPs of individual interface [32]. Among the multilayer structures, the metal-insulator-metal (MIM) configurations involving GSPs are superior to insulator-metal-insulator resonators in terms of high-quality factor and strong mode confinement [32, 33]. Here we will briefly discuss three typical GSP resonators with MIM configuration and present their scattering properties.

We begin the discussion of GSP resonator by revisiting the retardation-based resonant nanostructure realized by truncating the infinitely long MIM configuration, which consists of two gold strips separated by a silicon dioxide (SiO2) spacer (left panel of Figure 1A). Because of the terminations of the MIM structure, multiple reflections of GSP modes can be achieved, resulting in lateral standing-wave GSP modes, whose resonance positions can be described by a simple Fabry-Perot (FP) resonator formula [33, 35]:

\[
\frac{2\pi}{\lambda_0} n_{\text{GSP}} = m\pi - \phi,
\]

where \(\lambda_0\) is the wavelength in free space, \(w\) is the strip width, \(n_{\text{GSP}}\) is the real part of the effective mode index of GSP, \(m\) is an integer defining the order of the GSP mode, and \(\phi\) is an additional phase shift acquired upon reflection of the GSPs at the boundaries of the MIM configuration. By properly tuning the dimensions according to equation (1), GSP resonance at ~800 nm can be achieved with a gap size of \(d = 50\) nm, which shows a dominant scattering process, making this structure a good candidate for reflective metasurfaces (middle panel of Figure 1A) [34]. Additionally, the GSP mode at 800 nm shows a characteristic of magnetic dipole (MD) resonance with out-of-phase currents in the two metal strips and a maximum of the magnetic field in the gap (corresponding to a minimum of the electric field; right panel of Figure 1A) [36–38]. It is worth pointing that when the gap size is small, the GSP resonator is mainly an absorbing element because of the strong NFC between two metallic layers [39–41]. Once the gap size grows too large, the NFC becomes too weak to support GSP resonance, and the resonance undergoes a transition from a lateral MD resonance to a vertical Fabry-Perot resonance [42].

Besides the highly symmetric configuration, the GSP resonator can be constructed with metal strip placed close to an infinitely extended metal surface, which shows similar scattering resonances related to the GSP mode [43]. From Figure 1B, one can clearly see that the first-order GSP mode (i.e. MD resonance) has a red-shift of ~25 nm, which is ascribed to the less sharp termination of the resonator, allowing the GSP resonator to have a slightly larger effective width [34]. In addition, such resonator shows stronger scattering with a two-fold increase in the scattering cross section. One should note that such resonator can couple part of the scattered light into SPPs propagating along the metal-air interface.

As a third kind of GSP resonator, nanoscale truncation is carried out only for the top metal layer in a MIM structure, which can be termed as continuous-layer GSP resonator [44] (Figure 1C). In addition to the attractive properties of usual GSP resonators, this continuous-layer GSP resonator can be easily achieved by depositing gold in the resist profile with a single lithography step, thus relaxing the lift-off process. Compared with the second configuration, the GSP mode shows a strong red-shift
and is excited at $\lambda = 915$ nm, resulted from the removal of horizontal refractive index contrast between air and SiO$_2$ spacer [34]. Because of the simplified fabrication process, most of the GSMPs are constructed by the third kind of GSP resonators, in which the upper metal layer usually consists of periodic arranged antennas.

As a final comment, it ought to be mentioned that the GSP resonator could be well described by the coupled-mode theory [45–47]. Base on coupled-mode theory, Zhou’s group has established a generic phase diagram where the optical properties of GSP resonator are fully controlled by the intrinsic and radiative losses, thereby greatly facilitating the design of appropriate GSMPs with tailored functionalities [46].

### 2.2 Generalized Snell’s law

The abrupt phase shift, denoted as phase discontinuity, at the interface between two media was first introduced by Capasso’s group in 2011, which allows us to revisit the laws of reflection and refraction by applying Fermat’s principle, thereby realizing generalized Snell’s laws of reflection and refraction [2]. Figure 2 schematically illustrates the one-dimensional (1D) system used to derive the generalized Snell’s laws, where the interface between two
media is consisting of artificially metasurface in order to introduce a position-dependent phase shift $\Phi(x)$ in the light path. Considering an incident plane wave impinging at angle of $\theta$, relative to the surface normal of the metasurface-decorated interface, the generalized laws of reflection and refraction can be written as

$$\sin(\theta_i)n_i - \sin(\theta)n_t = \frac{l_0}{2\pi} \frac{d\Phi}{dx}, \quad (2)$$

$$\sin(\theta_i)n_i - \sin(\theta)n_t = \frac{l_0}{2\pi} \frac{d\Phi}{dx}, \quad (3)$$

where $n_i$ and $n_t$ are the refractive indexes of the two media, $\lambda_0$ is the wavelength in free space, and $\theta_i$ and $\theta$ are the reflected and refracted angles, respectively. From equations (2) and (3), it is evident that the reflection/refracted beam can have an arbitrary direction, provided that a suitable, constant, and nonzero phase gradient along the interface (d$\Phi$/dx) is introduced. In the case of d$\Phi$/dx = 0, we recover the usual laws of reflection of refraction that imply continuity of the in-plane wave vector.

To realize a metasurface with interface discontinuities, a V-shaped antenna, supporting symmetrical and antisymmetrical electric dipole resonance, was first proposed and experimentally demonstrated to control the reflection/refraction of linearly polarized (LP) light at middle-infrared range governed by the generalized Snell’s law [2]. Soon after, the idea was extended to the near-infrared regime, and the functionality was found to be wideband [48]. It should be noted that while great successes have already been achieved, there are two issues unsolved in the previous metasurfaces based on a V-shaped antenna [2, 48]. In these cases, the $2\pi$ phase control is achieved with reflected/transmitted light polarized orthogonally to the incident wave as the V-shaped antenna only supports electric dipole resonance and thus allows one to control the phase in a limited range, from 0 to $\pi$, because of its Lorentz-like polarizabilities [49, 50]. Such polarization conversion will cause inconvenience for some applications. Moreover, the theoretical upper bound of the coupling efficiency between the two polarizations for such single nonmagnetic metasurfaces is 25% in the limit of negligible absorption [51, 52]. To solve these two issues, one can place the metal antennas in close proximity to a back metal reflector separated by an optically thin dielectric spacer, making the GSPMs, which ensure an almost full control of the phase space together with a high efficiency reaching 100% for the lossless system for small angles [53].

One final point ought to be mentioned is that the generalized Snell’s law is not enough for realizing perfectly performing GSPMs because of significant parasitic reflections in undesired directions, especially for large incident angles [54]. To perfectly control reflection, one may use properly engineered bi-anisotropic meta-atoms [55, 56] or strong nonlocal response [57].

### 2.3 PB phase

In the previous examples, an abrupt phase discontinuity is introduced by varying meta-atoms’ geometric parameters; another completely different technique, known as PB phase or geometric phase, achieves a full phase control of the cross-polarized light by using anisotropic meta-atoms with identical geometry but spatially varying orientations for circularly polarized (CP) light [58, 59]. Here it should be emphasized that the recent development of metasurfaces based on PB phase has been largely following the early work on space-variant subwavelength gratings by Hasman and co-workers [60–63]. The concept behind PB phase is simple, which can be easily revealed by Jones matrix [64]. In general, the Jones matrix of an anisotropic GSP meta-atom that rotated an angle $\theta$ from the $x$-axis can be written as

$$M_x = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} M_{xx} & 0 \\ 0 & M_{yy} \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad (4)$$

where $M_{xx}$ and $M_{yy}$ are the reflection or transmission coefficients for LP light along the two primary axes of the anisotropic meta-atom, respectively, and $R(\theta)$ is the rotation matrix. When the incident light is CP, the reflected or transmitted light can be written as

$$M_0 \cdot E_0 = \frac{1}{2} (M_{xx} + M_{yy}) E_0^+ + \frac{1}{2} (M_{xx} - M_{yy}) e^{i2\theta} E_0^-, \quad (5)$$

where $E_0^+$ represents the incident left-handed CP (LCP) and right-handed CP (RCP) light, respectively. The first term in equation (5) represents CP light with the same helicity as the incident wave, while the second term stands for CP light with the opposite helicity that gains an additional PB phase of $\pm \theta$, whose sign depends on the helicity state. Thereby, a $2\pi$ phase coverage can be achieved if the meta-atom is rotated from 0 to 180$. In this sense, the PB phase can be regarded as dispersionless. However, the conversion efficiency may be strongly wavelength-dependent, as it relies on the ability of individual meta-atom to tune $M_{xx}$ and $M_{yy}$. It is clear that once the condition $M_{yy} = -M_{xx}$
is satisfied, the co-polarized light will vanish and the entire light is converted to the cross-polarization. As such, GSPMs play an important role in realizing highly efficient PB phase metasurface with the conversion efficiency up to 100% in the limit of negligible Ohmic loss [65].

3 Applications of GSPMs

GSPMs have been utilized in a wide range of applications, ranging from flat and compact versions of conventional optical elements with control over the amplitude, phase, and polarization of incident wave to dynamically reconfigurable GSPMs. Here we try to give an overview of the many, yet very different, applications.

3.1 Beam steering

As described in Section 2.2, beam steering is a fascinating functionality that can be realized by metasurfaces with constant phase gradient along the surface, leading to generalized Snell’s law of reflection and refraction that allows light to propagate in anomalous directions. Previous attempts to demonstrate beam steering typically relied on single-layered metasurfaces, which suffers from low efficiency [2, 48]. In contrast, GSPMs enabling full control of phase with high efficiency are an excellent candidate to realize highly efficient beam steering for both the LP light and CP light.

The first experimental demonstration of GSPM for beam steering is shown in Figure 3A, where H-shaped antennas are on top of a dielectric layer coated metal film, designed to anomalously reflect x-polarized incident wave in the microwave regime [53]. Through engineering the top H-shaped meta-atoms properly, a constant phase gradient is introduced along the x-axis, thereby resulting in diffraction into +1 order, verified by the angle-resolved far-field measurements (right panel of Figure 3A). As there is no polarization converted to the cross-polarized component and Ohmic loss is negligible, this GSPM shows close to 100% efficiency. Here it is worth noting that this type of GSPMs can be easily extended to near-infrared range with slightly reduced efficiency of ~80% because of increased Ohmic loss [69, 70].

Besides the high efficiency in controlling the reflection phase, GSPMs show even greater control of the reflected light by tailoring the reflection coefficients for two polarizations independently, thereby allowing for the design of birefringent GSPMs that achieve beam steering and splitting of individual LP light simultaneously [34, 66, 71–73]. As an example, Figure 3B shows a MIM unit cell for a birefringent GSPM working at 800 nm [66]. By properly selecting nanobricks with different widths, the incident light experiences an incremental (decremental) phase step within the supercell for the x-polarized (y-polarized) light, resulting in anomalous diffraction into +1 (−1) diffraction order (right panel of Figure 3B). The proof-of-concept experiment shows a reduced efficiency of ~50% compared with the theoretical efficiency around 80%, which is ascribed to the fabrication imperfections and uncertainties in the material properties. In principle, the polarization splitter functionality can be easily extended to visible by scaling the dimensions of the unit cell, which however requires small sizes and high accuracy of rectangular antennas. To relax these fabrication constraints, high-order GSP resonance was introduced to realize polarization beam splitter with excellent quality in terms of angular-independence and high polarization contrast [73].

Although the high-order GSPM could relax the difficulty in fabrication, it is rather challenging to experimentally realize metasurfaces with high-performance that operate in the visible range because of difficulties in fabricating several meta-atoms of subwavelength size with high uniformity. As an alternative, continuous rather than discrete phase delay using single spatial-variant gradient meta-atom is proposed to achieve beam steering [67, 74, 75]. Here we highlight a broadband and highly efficient anomalous meta-reflectarray across the visible frequency range by using trapezoid-shaped nanorod as the building block, which is quite easy to fabricate with high accuracy (Figure 3C) [67]. The broadband anomalous reflection and spectrum splitting have been experimentally confirmed at visible and near-infrared frequencies, shown in the right panel of Figure 3C. Specifically, most of the incoming power is reflected toward the anomalous direction (+1 order diffraction), while other diffraction modes are significantly suppressed. The average power ratio of +1 order diffraction to other diffractions for visible bandwidth from 500 to 800 nm is on the order of 106.

The resonant phase realized by tailoring the geometry dimensions of meta-atoms is intrinsically dispersive, and thus, the operating bandwidth is restricted. Distinct from the resonant phase, the PB phase is dispersionless in the sense that the induced phase on the cross-polarized light is based on identical meta-atom with spatially varying orientations as discussed in Section 2.3. Within the concept of GSPMs, Figure 3D depicts a highly efficient PB phase metasurface working at near-infrared regime, which consists of identical split rings with...
different opening orientations on the surface of a SiO$_2$-silver bilayer [68]. The right panel of Figure 3D shows the measured intensity of LCP and RCP components of two anomalous beams when the LP light impinges normally on the surface, confirming that the anomalous beam +1 is RCP and the anomalous beam −1 is LCP. Because of the
high polarization conversion ratio and broad bandwidth realized by the rotating split ring, this structure possesses significantly strong anomalous reflection beams (over 70% of incident light intensity) over a broad frequency range (1100–1750 nm). We note that a similar GSPM design has been realized in the microwave range with a record-high conversion efficiency of >90% [65].

3.2 Planar metalens

In conventional optical lenses, the wavefront engineering is typically based on the propagation phase gradually accumulated along the optical path. In this way, the dielectric lenses or parabolic reflectors with particular surface topography are bulky and curved components, thereby preventing the trend of very dense integration and miniaturization in photonics. Metasurfaces, on the other hand, enable an unprecedented opportunity to engineer the wavefronts at will, because of their planar profiles and relative ease of fabrication. To realize flat lenses, metasurfaces should impose the following parabolic phase profile to convert the incident planar wavefronts into spherical ones in transmission or reflection:

$$\phi(x, y) = \frac{2\pi}{\lambda_0}\left(\sqrt{x^2 + y^2 + f^2} - f\right),$$

where the metasurface is assumed to lie in the $xy$-plane, $f$ is the focal length, and $\lambda_0$ is the wavelength in free space. Here we review some of planar metalenses based on GSPMs, while also addressing the important topic of achromatic lenses. Starting with polarization-sensitive GSP metalens for 1D focusing [76, 77], Figure 4A conveniently illustrates the scanning electron microscope (SEM) image of the fabricated metalens, consisting of an array of gold nanobricks atop a thin dielectric spacer and continuous gold film. The dimensions of the meta-atoms are varied along the horizontal direction so as to achieve a parabolic phase profile. Here the continuously varying phase profile is approximated by a step-like function, where the phase takes on a finite number of values, thus making the metasurfaces consist of a series of aperiodic meta-atoms. The fabricated metalens can focus an LP (along the direction of nanobrick size variation) incident beam in the plane of its polarization with the focal length, which changes from $\sim$15 to 11 μm when tuning the light wavelength from 750 to 950 nm, respectively, verifying the broadband focusing ability. As shown in the lower panel of Figure 4A, the highest efficiency (27%) was measured at the wavelength of 800 nm. Even though the measured efficiency is considerably smaller than the calculated one because of various fabrication-related circumstances, it is substantially larger than the upper limit of ~10% estimated for V-antenna metasurfaces [81, 82]. We emphasize that polarization-independent focusing can be realized using symmetrical meta-atoms.

As stated in the previous section, birefringent metasurfaces have even greater capability of controlling the reflected light with different polarizations independently by engineering the reflection amplitudes and phases for two orthogonal polarizations simultaneously [34, 66, 71–73]. Therefore, by designing proper birefringent metasurfaces and imposing different parabolic phase profiles, a multifunctional GSP-based metalens (PSFML) has been experimentally demonstrated (Figure 4B), which can split orthogonal linear light polarizations and focus them into different focal spots [78]. The functionalities of polarization splitting and focusing are visualized in the right lower panel of Figure 4B, where it is readily seen that reflected light is focused in different positions for different polarizations. As can be seen, the PSFML deflects TM polarization to the right-hand side and TE polarization to the left-hand side from the origin, while two spots are symmetrically generated when two polarizations are involved. The proof-of-concept PSFML features high efficiency (up to ~65%) and polarization extinction ratio (up to ~30 dB), exhibiting broadband response in the near-infrared ranging from 750 to 950 nm.

Aside from the excellent focusing function, metalenses can provide scanning ability with controllable focus [79]. Figure 4C schematically shows the operating principle of the frequency-controlled focus scanning reflective metasurfaces, where the focal position is well controlled by varying the operating frequency for the plane waves incident at an angle of 45°. Specifically, the proposed flat metalens is successfully implemented by carefully controlling the dispersions of the elements with different physical dimensions, which are consisting ring and I-shaped resonators on top of a ground dielectric substrate. The frequency-controlled focus scanning ability has been numerically and experimentally confirmed for several wavelengths, with the lower panel of Figure 4C displaying the power intensity distribution along the x-axis at $z=40$ mm and $y=0$ mm. It can be seen that the measured results agree well with simulation results, which clearly demonstrate the focal-scanning ability of the flat metalens.

As conventional lenses suffer from severe chromatic aberrations because of the dispersion of the phase accumulated during light propagation, metalenses also face the same problem because of the intrinsic spectral
dispersion of the meta-atoms and the fact that phase gradient metasurfaces follow the diffraction theory [83]. As such, metalenses imply a change in focal length proportional to $\lambda^{-1}$, thus degrading the performance of an imaging system. Recently, several efforts have been dedicated to eliminate chromatic aberration of metalens by either dispersive phase compensation using multiple meta-atoms [84, 85] or birefringent unit cells with equal focal length for orthogonally polarized light at two different wavelengths [86]. However, the aforementioned metalenses can only compensate spectral dispersion at several discrete wavelengths, meaning that wavelengths of light deviating from the designed ones feature varying focal lengths. To achieve broadband achromatic meta-surface devices, Tsai’s group employed the integrated resonant unit element in metasurfaces with smooth and linear phase dispersion combined with the PB phase [80]. The proof-of-concept broadband achromatic metalens is capable of focusing light with arbitrary wavelength at the same focal plane with the efficiency on the order of $\sim12\%$ over a continuous wavelength region from 1200 to 1680 nm for CP incidences in a reflection scheme (Figure 4D). Additionally, this approach is not limited to achromatic lens and thus can be easily used to realize
various flat achromatic devices, such as achromatic gradient metasurfaces. Freshly, a polarization-insensitive achromatic metalens was demonstrated by Capasso’s group to suppress the chromatic effect over a 60-nm bandwidth in the visible for LP incidences, which consists of titanium dioxide (TiO$_2$) nanopillar array on a silica-coated silver substrate. However, it is worth pointing out that the phase gradient is mainly due to the waveguide mode supported by the top TiO$_2$ nanopillar. In this way, the structure cannot be regarded as GSPM [87].

3.3 GSPM-based hologram

Conventional holograms, based on the phase accumulation along the optical path, suffer from wavelength-scale thickness and large pixel size, which is usually tens of wavelengths for visible light, resulting in severe issues of high-order diffraction and twin image. In contrast, arbitrary amplitude, phase, and polarization can be locally engineered by introducing resonant meta-atoms and tailoring their lateral dimensions or orientations while the thickness is at the subwavelength scale. As such, metasurfaces are a good platform to realize all types of computer-generated holograms [24]. Figure 5 shows several typical examples of GSPM-based holograms with excellent performance.

To reconstruct a holographic image, a $2\pi$-phase modulation with quasi-uniform amplitude distribution is usually required, which can be achieved using resonant phase through deliberately optimizing the structure geometry for LP light at the design wavelength. Figure 5A schematically illustrates the first demonstration of GSPM-based hologram that generates reflective hologram at visible wavelengths [88]. Particularly, by utilizing the polarization-sensitive response of the gold cross nanoantennas, two holograms with different patterns are simultaneously encoded on the same GSPM, and each image is reconstructed with an LP light with orthogonal polarizations, which works as a metasurface-based polarization-multiplexing technique. The performance of polarization controlled dual images was experimentally verified, where the projected patterns transitioned from NTU to RCAS once the input polarization is changed from $x$- to $y$-direction. The measured efficiency at $\lambda = 780$ nm reaches 18% at 15° incidence and drops down to 8.5% when the angle is increased to 45°, which is greater than the existing meta-holograms with fishnet [92] and complementary V-shaped

![Figure 5: GSPM-based holograms.](image)

(A) Polarization-multiplexing holograms based on meta-reflectarrays for linear polarization states. Reprinted from Ref. [88]. (B) Sketch of multicolor holograms using aluminum nanorods as the basic building block. Reproduced from Ref. [89]. (C) Highly-efficient reflective holograms employing PB phase and MIM configuration for circular polarization. The inset shows a SEM image of a part of the fabricated sample. Reprinted from Ref. [90]. (D) Diagram of a helicity-multiplexed hologram realized by a reflective PB-phase metasurface. The inset shows a SEM image of a part of the fabricated sample. Reprinted from Ref. [91].
structures [93]. We note that the efficiency of GSPM-based hologram can be greatly increased to around 50% at a near-infrared wavelength of 1550 nm, resulting from the decreased metal loss [94].

Besides polarization-multiplexing, wavelength-multiplexed metahologram operating under LP light illumination was also demonstrated, where different holographic images were constructed by addressing the incident wavelength [89]. With proper design of the aluminum nanorods with different sizes, resonances with narrow bandwidths are realized, thereby allowing for implementation of the multicolor metahologram with two-phase modulation at three primary colors (Figure 5B). However, it should be noted that the efficiency of wavelength-multiplexing metaholograms is inevitably low, less than 1%, ascribed to the trade-off between wavelength and phase control to minimize the crosstalk between different color images by spectrally separating the multiple resonances [89].

In the previous GSPM-based holograms, phase modulation is achieved by geometry-dependent resonance, thereby limiting the operating bandwidth because of the intrinsic dispersive resonance of meta-atoms. As an alternative approach to metaholograms, PB phase holograms gain the required phase on the cross-polarized light with identical geometric parameters but spatially varying orientations, thereby releasing the process of design and fabrication. However, it is worth noting that although PB phase modulation is dispersionless, the working bandwidth of PB phase metaholograms is still limited, which is in line with Section 2.3. Because of the linear relationship between the PB phase and the orientated angle of the individual meta-atom, the design of multilevel-phase holograms becomes much simpler. As an example of GSPM holograms based on PB phase, Figure 5C schematically illustrates the holographic image reconstruction using a GSPM that introduced 16-level geometric phases [90], where the MIM structure was optimized as a broadband half-wave plate maintaining high polarization conversion. Remarkably, this metahologram has a diffraction efficiency of 80% at 825 nm resulting from the antenna-orientation controlled PB profile combined with the GSPM for high polarization conversion. Additionally, the dispersionless nature of the PB phase GSPM results in broadband operation between 630 and 1050 nm with a high window efficiency larger than 50%.

Similar to the resonant phase holograms, holography multiplexing, such as spin-multiplexing, can be easily accomplished with PB phase metasurfaces. Figure 5D illustrates a good example of helicity multiplexing with two sets of PB phase GSPMs, which functions as synthetic holograms for normal incident CP light in reflection [91]. Two symmetrically distributed off-axis images are simultaneously reconstructed in one identical hologram by controlling the helicity of the incident light. Moreover, the positions of the symmetrically reconstructed holographic images are exchanged once the spin of the incident light is altered. The experimentally measured conversion efficiency is higher than 40% over a broad wavelength band ranging from 620 to 1020 nm, with the maximum conversion efficiency reaching 59.2% at $\lambda = 860$ nm. Such high image quality and wide bandwidth are ascribed to the simple structure geometry and the geometric nature of the PB phase, which is more robust against fabrication tolerances and variation of material properties.

### 3.4 Ultrathin absorbers

Besides the aforementioned functionalities, enhancing absorption of light in a controllable manner is also of central importance to numerous applications such as photovoltaics [31], photodetectors [95, 96], thermal emitters [97], and biomedicine [98]. Whereas bulky plasmonic materials, for example, noble metals, already exhibit substantial absorption, coexisted strong reflection and poorly-defined spectra features severely limit their further applicability and practicality. To resolve this issue, a particular branch of metasurfaces, perfect absorbers with near-unity absorption and designer spectra, has emerged and garnered considerable interest in recent years. In this context, GSP resonators with intriguing characteristics provide novel possibilities to enhance light absorption processes. To achieve perfect absorption, a system needs to prohibit all transmission and reflection channels at a certain frequency range. In view of this, optically thick continuous films are generally adopted as bottom back reflectors to block all transmission, making GSPMs as single-port reflective systems. By rationally designing topmost metallic nanostructures and carefully choosing spacer thickness, reflection can be eliminated via the generation of strongly confined GSP resonances. This exotic effect has triggered intense research efforts to realize GSP absorbers with various functions, aiming at many different uses in a diverse set of scenarios [99–101]. Here we divide GSP absorbers into two main categories based on their different operating bandwidth, i.e. narrowband absorbers and broadband absorbers.

#### 3.4.1 Narrowband absorbers

Figure 6A illustrates a representative narrowband perfect absorber using GSP resonators as basic building
Figure 6: Ultrathin absorbers.
(A) Left panel: schematic of the perfect absorber structure and SEM image of the fabricated sample. Right panel: experimental tuning of the reflectance and absorbance spectra by changing the dielectric environment that is adjacent to the gold disks from air to water. Reproduced from Ref. [39]. (B) Left panel: schematic diagrams for the photothermal reshaping of gold nanoparticles in a plasmonic absorber. Right panel: SEM image of a region of the irradiated sample with both unmelted (upper) and melted (lower) gold nanoparticles. Reprinted from Ref. [102]. (C) Upper panel: schematic representation of the quadratic unit cell containing four gold nanoparticles with different diameters. Lower panel: reflection spectra for unpolarized (solid red curve), x-polarized (dashed blue curve), and y-polarized (black dash-dotted curve) light. The inset shows the SEM image of the sample and the microscope image of the absorber for unpolarized light. Reprinted from Ref. [103]. (D) Upper panel: schematic of the broadband absorber made up of highly lossy metal. Lower panel: simulated (red dashed curve) and measured (black solid curve) absorption spectra for unpolarized light. The inset shows the SEM image of the sample. Reprinted from Ref. [104].
blocks [39]. The top-layer gold disks were laid out with subwavelength periodicity, in a similar manner to other GSPMs. However, unlike reflective GSPMs, GSP absorbers usually require a much thinner dielectric spacer to ensure strong interactions between the top nanoparticles and the bottom reflector, forming a closely coupled antisymmetric current distribution [40, 41]. Such a field distribution resembles an electric dipole near a mirror, giving rise to a mirrored dipole approximately in antiphase and thereby canceling out far-field radiation of the system, which can be also regarded as MD resonance. A maximum storage of near-field energy is thus generated, and 99% absorption can be achieved in the near-infrared range at normal incidence, as shown in the right panel of Figure 6A. Because of the rotational symmetry of the design and its ultrathin thickness, such a high absorption is polarization independent and robust against a wide range of incident angles. Moreover, when the GSP absorber is covered with different dielectric surroundings, the spectra show evident shifts or noticeable changes in the reflectivity/absorptivity at a fixed frequency. These interesting features further provide the feasibility of GSP absorbers working as plasmonic sensors.

After absorbing light, the confined EM field inside GSP resonators then dissipates into thermal energy, resulting in localized heating response and associated photothermal effects. With nearly full absorption, GSP absorbers manifest a strikingly enhanced photothermal process that is not reachable with natural materials or conventional plasmonic nanoparticles. Even with a moderate incident power (i.e. 2.3 mW), GSP absorbers can generate an ultra-high temperature elevation up to 900 K [102]. As a consequence, the top metallic nanoparticles would undergo melting and reshaping by light irradiation, as displayed in Figure 6B. Further transient analysis demonstrated that such a remarkable temperature increase occurs in just a few nanoseconds [105]. The strong photothermal effects and associated fast response make GSP absorbers a promising candidate for many temperature-related applications including biotechnology [106], modulation [107–109], laser-induced particle transfer [110], and the rapidly growing field of plasmonic color laser printing [111].

### 3.4.2 Broadband absorbers

So far, the foregoing absorbers featuring single GSP resonance all possess narrowband response, which is desirable for certain practical scenes such as plasmonic sensing [39, 112]. However, for many other applications, especially photovoltaics and energy harvesting, broadband absorption is usually highly in demand. The most intuitive way to realize broadband GSP absorbers is to utilize GSP resonators with multiresonant properties. For example, by combining distinct yet overlapped GSP resonances within one design, a broadband response can be obtained, as shown in Figure 6C. When four differently sized GSP resonators and nine identical nanoparticles constitute a single unit cell, average absorption of ~90% for unpolarized light can be achieved in the entire visible spectrum (400–750 nm) [103]. This simple, yet powerful design principle can also be leveraged to other configurations and other material platforms [113]. Indeed, metallic nanostructures on the top of GSP resonators can be further tailored to introduce additional functionalities such as polarization dependence and dual- or multi-band absorption [114–119]. We also note that chemically synthesized nanoparticles can serve as the top metallic nanostructures [120–122]. By doing this, GSP absorbers can overcome the spatial limitation of top-down lithography techniques, realizing perfect absorption over truly macroscopic areas. In addition, by controlling the size distribution and stacking layers of self-assembled nanoparticles, spectral response of GSP absorbers can be further engineered toward different practical applications.

Figure 6D illustrate another route to attain broadband absorption, i.e. the exploitation of highly lossy metals instead of commonly used coinage metals [104]. Periodic titanium nanodisks were placed at the top of a SiO₂ spacer layer and a continuous gold film. The pronounced inherent loss in titanium decreases the quality factor of GSP resonances and consequently broadens the absorption bandwidth. The accompanied propagating SPP along the SiO₂-gold interface provides an extra channel for light dissipation. Absorption exceeding 90% thus was experimentally observed over a wide spectrum ranging from 900 to 1825 nm. Other highly lossy materials such as tungsten and nickel also hold the promise for broadband absorption [123–125].

### 3.5 High-resolution color printing

The printing techniques in industry, such as inkjet and laserjet methods, use micro-sized ink spots and result in limited dots per inch (DPI), which is usually less than 10,000. Worst still, the positioning of multiple color pigments is inaccurate. To increase the resolution of color printing, significant attention has been drawn to designing proper micro/nanostructures for structural colors, inspired by the natural colors. In this sense, metasurfaces are a good platform to generate vivid colors through...
a strong interaction between light and subwavelength nanostructures [126–129]. In particular, GSPMs, which enable deep-subwavelength light confinement [30], can print subwavelength color elements beating the diffraction limit as MIM structures can be tailored for efficient light absorption by properly changing the optical constant or the thickness of the middle spacer layer that separates the two metal films, as discussed in the previous section. Additionally, the operating wavelength of MIM structures can be easily tuned with the different dimensions, leading to flexible spatial and spectral localization of light. In this section, we will discuss recent advances in GSPM-based high-resolution color printing, and several kinds of representative GSPMs that have excellent quality or special functionalities of structured colors are presented.

Figure 7A shows an example of polarization-insensitive GSP color printing with MIM structure containing gold nanodisk arrays hovering on a continuous gold film with a thin dielectric layer in between [130]. When the size of the top nanodisks and the thickness of the dielectric layer are changed, the reflected spectra shift accordingly, resulting in tunable color. With properly selected disk sizes, an optical image of the University of Southern Denmark logo was produced, revealing a high-quality color print in terms of high contrast, high color uniformity, and high reproduction fidelity. Remarkably, this GSPM color printing is insensitive to the incident angle, superior to other structured color that is angular-dependent. Additionally, a cover of transparent dielectric overlay on top the fabricated GSPM would not affect the color too much, as shown in the right panel of Figure 7A. One can clearly see that the overall color image remains except for a slight contrast variation, ascribed to the nature of GSP resonance that entirely concentrates the light within the MIM structure itself. Therefore, this GSPM color device can be protected with a transparent dielectric overlay without destroying the colors. Similarly, aluminum-based GSPMs have also been considered for full-color subwavelength printing [132].

To tackle the challenge of highly saturated and bright color generation with plasmonic structures and maintain the pixel design functionality, tandem nanodisk arrays were demonstrated to achieve full colors [131] (Figure 7B). Here the MIM sandwich nanodisks are employed, which support enhanced in-phase electric dipole mode showing blue shift with respect to a single metal disk. By combining this dipole mode with the Wood’s anomaly in periodic structures, a high narrow peak in reflection and a deep valley in transmission is produced, thus providing vivid red-green-blue colors in reflection and cyan-magenta-yellow colors in transmission. The right panel of Figure 7B displays the experimental reflection and transmission colors for fabricated capital letters in serif font Georgia captured by optical microscope, respectively, where the colors are kept complementary in reflection and transmission and the full details of all letters are clearly reproduced.

The previous GSPM-based color patterns are relying on electron-beam lithography, which is expensive and not suitable for post-processing customization. Worst still, the color patterns need to be redesigned and thus are not scalable when the electron-beam lithography is conducted. More recently, printing new colors on a prefabricated GSPM color substrate using laser post-writing was successfully demonstrated (Figure 7C) [111]. Laser pulses induce transient local heat generation that results in melting and reshaping of the nanoimprinted gap-plasmon nanostructures [102, 105]. Depending on the laser pulse energy
density, different surface morphologies can be created, corresponding to different gap-plasmon resonances and consequently different color modulation. Based on this technique, all primary colors can be printed with a speed of 1 ns per pixel, resolution up to 127,000 DPI, and power consumption down to 0.3 nJ per pixel. As a final comment, it should be mentioned that although laser-induced morphology changes are irreversible, the structures can be repeatedly rewritten to exhibit new color resonances associated with morphologies of lower surface energy.

### 3.6 Polarimeter control and detection

#### 3.6.1 Polarimeter control

Polarization is an intrinsic property of EM waves. The capability of controlling the polarization states is of vital importance in a variety of practical applications. Typically, the conversion of polarization states is achieved by utilizing birefringent crystals, where the phase retardation between two orthogonally polarized wave components is gradually accumulated during light propagation. As a result, the polarization control components are usually bulky and suffer from specific thickness limitations, going against the general trend of integration and miniaturization in photonics, which in turn boosts the development of metasurface-based waveplates operating at different wavelengths. In this section, we review some examples of the GSPM waveplates with excellent performance (Figure 8).

Generally, GSPMs featuring anisotropic optical responses for two orthogonally polarized wave components, i.e. $E_x$ and $E_y$, can be applied to the realization of ultrathin metasurfaces based quarter-wave plate [137–140] or half-wave plate (HWP) [133, 135, 141–146] when
the reflection amplitudes $R_{xx}$ and $R_{yy}$ are equal, and the phase difference $\Delta \Phi = \Phi_{xx} - \Phi_{yy}$ is $\pi/2$ or $\pi$, respectively. For instance, Anders et al. demonstrated that the MIM configurations in which the top metal layer consists of a periodic arrangement of nanobricks whose main axes function as orthogonal detuned scatterers, thus supporting detuned GSP resonances, can be designed to function as reflective broadband half-wave plates [133] (Figure 8A). By properly tuning the dimensions of nanobricks, two resonant dips can be found under the illumination of horizontally and vertically LP light, corresponding to the GSP resonances along the long- and short-axes, respectively. Therefore, the reflection efficiencies between the $x$- and $y$-polarized light exhibited nearly equal reflection amplitude ratio $R_{xx}/R_{yy}$ accompanied by a phase difference close to 180° in the off-resonance regime between two GSP resonances, leading to broadband linear-to-linear polarization conversion, as shown in the left lower panel of Figure 8A. From the measured reflectivity at different analyzer angles (right lower panel of Figure 8A), it is clearly demonstrated that the reflected light is rotated by 90° with the maximum reflection reaching ~50% at the analyzer angle 130° in the wavelength range of 740–880 nm. The broadband polarization conversion can also be interpreted in the general mechanism of dispersion engineering by integrating a metallic metasurface possessing a strong yet dispersive interaction with light, with a dielectric interlayer, which has accumulative yet broadband interaction with light [140]. In this way, the intrinsic dispersion of the metallic structures is perfectly canceled out by the thickness-dependent dispersion of the dielectric spacing layer, resulting in dispersion-free metasurface and thus wideband linear-to-linear or linear-to-circular polarization conversion.

Because of increased Ohmic loss, the efficiencies are slightly reduced when GSPMs are extended to near-infrared and visible regimes. To increase the efficiency, a hybrid GSPM was demonstrated to create broadband HWP [134]. The metasurface in Figure 8B consists of an array of 45°-rotated silicon bricks, 200-nm PMMA spacer, and a sliver back-reflector, thereby considerably reducing the level of absorption within the metasurface in near-infrared range. The experimental measurements demonstrate that the cross-polarized reflection $|r_{xy}|^2$ remains above 97% while the co-polarized reflection $|r_{xx}|^2$ is greatly suppressed from 1420 to 1620 nm, in good agreement with the simulations. Additionally, the measured polarization conversion ratio (PCR), defined as $|r_{xy}|^2/(|r_{xx}|^2+|r_{yy}|^2)$, remains above 98% over this bandwidth across a 200-nm bandwidth. Furthermore, by varying the meta-atom dimensions, the $2\pi$ phase coverage of the reflected cross-polarized light has been achieved while maintaining high conversion efficiency, opening the door to ultracompact reflective phase plates for vortex beam generation.

Though the GSPM-based wave plates can achieve excellent polarization conversion over a wide wavelength spectrum, the co-polarized light cannot be completely eliminated. As such, there will always be co-polarized reflected light remaining in space, merging together with the converted cross-polarized light and affecting the performance. To separate the co-polarized and cross-polarized reflection spatially, a background-free GSPM-based HWP was demonstrated by integrating several HWPs into a supercell to create a linear phase gradient for the cross-polarized reflected wavefront, resulting in anomalous reflection and routing the cross-polarized reflected light away (Figure 8C) [135]. Specifically, four antennas are introduced by numerical simulations to form nearly $2\pi$ phase shift coverage for the cross-polarized reflected light with a phase increment of about $\pi/2$ between adjacent elements, and each antenna can serve as a HWP with high PCR. As shown in the right panel of Figure 8C, the cross-polarized reflection power (red curve) coincides with the total reflection power (blue curve), and the co-polarized reflection power (black curve) is approaching zero, confirming the required performance. In addition, the ratio between the power of the desired and orthogonal polarizations is above 20, and the integrated PCR exceeds 95%.

In contrast to the aforementioned ultrathin wave plates based on resonant phases, the PB phase can be used to realize the broadband wave plate [136, 147, 148]. Here we would like to highlight a GSPM polarization generator, which is capable of producing reflected light beams of any polarizations from an LP light source (left panel of Figure 8D) [136]. The metasurface chip is composed of six types of supercells, in which each supercell is made up of two subunits with different offsets. Each subunit features a constant phase gradient by rotating the orientations of aluminum nanoantennas gradually with a constant step size of $\theta$ from 0 to $\pi$. Upon the incidence of LP light, two CP reflection beams with opposite handedness are redirected into the ±1 diffraction orders. If two supercells with different offsets and opposite nanoantenna orientations are introduced, different LP states could be produced. From the right panel of Figure 8D, it is clearly seen that six light beams with different polarizations are achieved in different reflection angles simultaneously, including four different linear polarizations and two circular polarizations.
3.6.2 Polarimeter detection

The rapid development of metasurface-based polarization conversion has motivated the implementation of polarimeters, which enables direct detection of polarization states. As the polarization state of light is uncorrected with its intensity and wavelength, it is intrinsically difficult to be measured. Typically, polarization states are probed by using a set of discrete polarizers and wave plates consecutively placed in the light path in front of a detector. By measuring the transmitted intensity through these components for polarization control, the Stokes parameters that uniquely define the polarization state can thus be determined. In this way, polarimeters based on conventional optical components are bulky and complicated, not compatible with the integrated photonic system. In recent years, polarization detection using GSPMs has gained much attention because of their ease of fabrication, compact footprint, and design flexibility [149–151]. For example, photonic spin Hall effect has been utilized to implement chiroptical spectroscopy based on PB phase GSPM, which can determine certain parts of the polarization state, i.e., the degree of circular polarization [149]. Such chiroptical spectroscopy can spatially separate the LCP and RCP light, and the scattering intensity ratio of these two CP states can be used to retrieve the ellipticity and handedness of the incident light. However, this type of chiroptical spectroscopy has only two channels and cannot fully characterize the polarization state.

Recently, on-chip GSPM polarimeter has been proposed to simultaneously determine all polarization states by a metagrating [150]. As shown in Figure 9A, the metagrating consists of three interwoven GSPMs in which each functions as a polarization beam-splitter for a certain polarization basis. Once an arbitrary polarized incident beam impinges on the metasurface, the beam will be diffracted spatially and routed in six predesigned directions. By conducting parallel measurements of the six diffraction intensities, the Stokes parameters can be quickly retrieved, thereby allowing one to analyze an arbitrary polarization state. At the design wavelength of 800 nm, the experimental diffraction contrast obtained by averaging three successive measurements agrees well with the input normalized Stokes parameters, repeating well the corresponding diffraction contrast being around ~0.1 (right panel of Figure 9A). Additionally, the designed GSPM could operate well in the wavelength range of 750–850 nm. Following this concept, a waveguide-based polarimeter was proposed based on GSPM, where the polarization state can be directly determined by calculating the coupling efficiencies of launched waveguide modes in six different directions [151].

To further extend the functionality of polarimeter, segmented and interleaved GSPMs have been proposed and demonstrated to realize spectropolarimeters, which enable simultaneous characterization of the polarization state and spectrum of incident light [152–154]. Starting with segmented GSPM spectropolarimeter, Figure 9B schematically illustrates the principle of a spectropolarimeter metadevice composed of six GSPMs arranged in a 2×3 array corresponding to horizontal (0°), vertical (90°), ±45°, RCP and LCP analyzers, which steers different polarization and spectrum components into distinct positions [152]. Once a probe beam with a certain polarization state is incident on the metadevice, six intensity peaks are generated in the far-field. After careful calibration, the polarization detection at a certain wavelength is determined. Additionally, the measured angular dispersion for the LCP and RCP channels are 0.053°/nm and 0.024°/nm, respectively (right panel of Figure 9B), proving the potential of spectral measurement. Later on, an interleaved GSPM spectropolarimeter was demonstrated by randomly interspersing three linear phase profiles associated with different meta-atom subarrays (Figure 9C) [153]. As shown in the middle panel of Figure 9C, the measured Stokes parameters are in good correspondence with the calculated ones on the Poincaré sphere for an arbitrary incident beam at the design wavelength of 760 nm, revealing the excellent capability of polarization probe. Furthermore, the interleaved spectropolarimeter has good spectral resolving power (Δ/Δλ), which was measured to be 13 when the diameter is only 50 μm.

Although these two spectropolarimeters show good performance, they are very sensitive to the size of incident beam, and the calibration is always needed. To increase the robustness of polarization and spectrum detection, a self-calibrating segmented GSPM spectropolarimeter, featuring the beam-size-invariant nature, has been demonstrated [154]. Figure 9D shows the GSPM spectropolarimeter with center-symmetrical configuration, which consists of three GSPMs that occupy 120° circular sectors each. Upon the excitation of a normally incident beam with arbitrary polarization state, six diffraction spots are generated in the predesigned directions, whose polar angles are proportional to the wavelength while the contrasts in the corresponding diffraction intensities provide a direct measure of the incident polarization state. Here it should be pointed out that uneven illumination of the three GSPMs caused by any misalignment will not affect the final result and no calibration is needed at all as the retrieved Stokes parameters are only associated with the
Figure 9: Polarization detection.
(A) Left panel: illustration of the metagrating’s principle. Right panel: measured diffraction contrasts (filled circles) for polarization states along the main axes of the Poincaré sphere at 800 nm. Reprinted from Ref. [150].
(B) Left panel: illustration of the integrated GSPM device. Right panel: theoretical predictions and experimentally observed spectral dispersion for circular detection channels. Reprinted from Ref. [152].
(C) Left panel: schematic setup of the spectropolarimeter. Middle panel: predicted (red dashed curve) and measured (blue circles) polarization states. Right panel: measured far-field intensities for elliptical polarization at two spectral lines (i.e. 740 and 780 nm) and the corresponding resolving power of the 50-μm-diameter device. Reprinted from Ref. [153].
(D) Left panel: illustration of the GSPM-based beam-size-invariant spectropolarimeter. Right panel: normalized measured far-field intensity profile for different wavelengths of the |x⟩ channel. Reprinted from Ref. [154].
relative diffraction contrasts for three polarization bases. The proof-of-concept 96-μm-diameter spectropolarimeter operating in the wavelength range of 750–950 nm exhibits excellent polarization sensitivity with beam-size-invariant property. Additionally, the experimentally measured spectral resolving power is estimated to be ~15.2 (right panel of Figure 9D).

3.7 Surface wave couplers

In addition to manipulating the propagating waves (PWs) in free space, another exciting feature of GSPMs is the capability of efficiently launching surface waves (SWs), such as SPPs and spoof SPPs in low-frequency range, in which arbitrary in-plane wave vector can be designed to match with the wave vector of SWs, distinct from the prism or grating couplers based on resonant coupling between PW and SW [155]. In this way, a nearly perfect PW-SW unidirectional conversion for any incidence angle larger than a critical value can be achieved [53, 156]. However, the generated SW is not the eigenmode of the GSPM and more likely to scatter out because of the spatial inhomogeneity of the GSPM. Additionally, the conversion efficiency decreases significantly with the increased size of incident beam, arising from the significant scattering caused by inter-supercell discontinuities [157]. To solve these issues, a new SPP meta-coupler has been demonstrated, which consists of a transparent gradient metasurface placed at a certain distance above the target plasmonic metal [158]. As shown in Figure 10A, the incident wave is first converted into a driven SW bound on the metasurface and then resonantly coupled to the eigenmode (i.e. SPP wave in optical frequencies) on the plasmonic metal. Based on this new configuration, a theoretical efficiency up to 94% has been predicted by model calculations, eliminating the nonnegligible issues that severely affect the coupling efficiency [157]. As a practical realization, a realistic device operating in the microwave regime has been fabricated, which exhibits a spoof-SPP conversion efficiency of ~73% in both near-field and far-field experiments (right panel of Figure 10A). The measured conversion efficiency is much higher than those of all other available devices operating in this frequency domain.

Though the aforementioned GSPM couplers show good performance, the direction of launched SWs is predefined without any tunability. Very recently, unidirectional polarization-controlled SPP coupling has been demonstrated using GSPMs, so that the direction of SPP excitation was switched by the polarization of an LP light [159] or the handedness of a CP incident beam [161, 162]. Here we would like to highlight an efficient unidirectional polarization-controlled SPP coupler by using GSPMs operating at telecom wavelengths [159]. Specifically, arrays of GSP resonators that would produce two independent orthogonal reflection phase gradients in two respective linear polarizations have been proposed, where basic GSP unit cell consists of a gold nanobrick atop a 50-nm-thick dielectric spacer layers covering 80-nm-thick gold films supported by glass substrates (Figure 10B). Therefore, the incident radiation (with arbitrary polarization) can efficiently (up to 40%) be converted into SPPs propagating in orthogonal directions dictated by the phase gradients. The right panel of Figure 10B displays the recorded images from leakage radiation microscopy for a free-space wavelength of 1500 nm, featuring the coupling efficiency C of ~25% for either of two linear polarizations and the directivity D of SPP excitation exceeding 100.

To further extend the functionalities of GSPM-based SW coupler, a bifunctional GSPM for visible light has been proposed, which is capable of efficiently converting x-polarized radiation into SPPs and anomalously steering the reflected light for y-polarization simultaneously (Figure 10C) [160]. Here two different linear phase gradients, ζx and ζy, are introduced along the x-direction in respective linear polarizations, distinct from the previous polarization-controlled GSPM coupler processing orthogonal phase gradients [159]. Specifically, the reflection phase gradient for the x-polarization is equal to the wave-vector of excited SPPs bounded on the air-dielectric-metal interface (i.e. ζx = ksp), while the phase-gradient is smaller than the wave-vector of propagating light in free space when the incident light is y-polarized (i.e. ζy < k). Using an x-polarized Gaussian beam optimally positioned to maximize the coupling efficiency of the right-propagating SPPs, unidirectional excitation of SPPs over a wide spectrum range was observed. Middle panel of Figure 10C shows the measured coupling efficiencies over the wavelength range from 600 to 650 nm, where an averaged coupling efficiency Cx of above 25% has been achieved, and the measured extinction ratio is larger than 20 dB. Apart from the unidirectional SPP excitation under x-polarization, the designed metasurface can function as a broadband beam steerer for the y-polarized light. From the right panel of Figure 10C, one can clearly see that most of the reflected light is contained within the ±1 diffraction order in a broad wavelength regime, verifying the broadband steering for the y-polarized light, albeit with some discrepancies compared with the simulation results regarding the efficiencies, especially at short-wavelength range.
3.8 Dynamically reconfigurable GSPMs

The aforementioned unprecedented phenomena and applications of GSPMs are mainly relying on their passive properties. An as-yet largely unexplored milestone in the field of metasurfaces is to achieve a dynamically reconfigurable metasurfaces, which can extend their exotic passive properties and enables dynamic wavefront control over an ultrathin surface, such as active spatial light modulation, dynamic beam steering, and tunable ultrathin lenses [22, 163].

Reconfigurable GSPMs can be realized by altering the shape of individual meta-atoms, or by manipulating the near-field interactions between them, which can be accomplished through microelectromechanical systems [164, 165]. For example, a voltage-tunable Fano metamaterial absorber, which simultaneously supports thermomechanically coupled optical and mechanical resonances, has been utilized to control mechanical damping with light (Figure 11A) [164]. Based on the plasmo-mechanical parametric gain mechanism, optically pumped coherent mechanical oscillations with a very low threshold have been experimentally demonstrated over a ~4-THz bandwidth. In addition, optical damping of mechanical resonance was also achieved through the reverse effect.

Besides the structure aspect, an alternative strategy to realize reconfigurable GSPMs is to incorporate active functional materials into metasurfaces that provide tunable optical constants through voltage bias, thermal heating, optical pump, and magnetic field. In particular, owing to the excellent mechanical and optoelectronic properties and the largely tunable carrier density, graphene
Figure 11: Dynamically reconfigurable metasurfaces.
(A) Left panel: schematic of the metamaterial absorber with a mechanically compliant bilayer (gold/SiN) membrane component decorated with a nanoantenna array. Right panel: plasmomechanical damping and the related mechanical resonance spectra. Reprinted from Ref. [164].
(B) Left panel: a schematic of the graphene GSPM, which consists of an array of aluminum mesas and a continuous aluminum film (yellow) separated by an SU8 spacer (blue). Right panel: gate-dependent reflection phase spectra of two different devices. Insets: the elements of devices A and B. Reprinted from Ref. [47].
(C) Left panel: illustration of the metasurface device consisting of a sandwich system with a patterned mesh top gold layer, an active VO2 layer, and an optically thick gold ground plane. Right panel: Measured reflection spectra for various intensities of electrical current applied. Reprinted from Ref. [166].
(D) Left panel: schematic of the gate-tunable ITO-integrated GSPM. Voltage is applied between the antenna array and the bottom gold resulting in the formation of charge accumulation at the Al2O3/ITO interface. Right panel: simulated and experimental measured far-field intensity profiles of the diffracted beam versus applied bias. Reprinted from Ref. [167].
processes largely tunable optical conductivity in the mid-infrared and THz frequency ranges once the doping is adjusted through changing the bias voltage, resulting in the ultimate reconfigurability and multiple functionalities with graphene GSPMs [47, 168–170]. Figure 11B schematically shows the wide-range phase modulator, which is consisting of magnetic resonators and gate-controlled graphene [47]. A variation in the graphene resistance, associated with the gate, drastically changes the resonance behavior of the proposed metasurface, namely, driving an underdamped resonator to an overdamped resonator, leading to a phase modulation of ±180°. It should be noted that the maximum phase modulation achieved using the present scheme is not limited to 180°. A much larger maximum phase modulation can be achieved using two slightly different graphene metasurfaces gated independently (right panel of Figure 11B). The large phase modulation is ascribed to the fact that gate doping modulates the phase in the range from approximately −180° to 0° at frequencies immediately above the resonance frequency in device A, while simultaneously suppressing the phase in the range from approximately 180° to 0° at frequencies slightly below the resonance in device B. Consequently, an extremely large phase modulation range of 243° at 0.48 THz has been achieved. The combination of graphene and GSPMs generates a versatile new platform for applications in the mid-infrared and THz frequency regime. Here it should be emphasized that highly doped graphene sheet or graphene nanostructures are needed if one wants to realize tunable graphene metasurfaces in the near-infrared and even the visible range [171].

As alternative functional materials with excellent reconfigurability, phase-change materials has been extensively utilized in optical data-storage systems and photonic devices because of their outstanding switchable optical properties by thermal, laser, or electrical current pulses with controlled duration and intensity [166, 172–179]. On the transformation from the amorphous to crystalline state, their optical properties change drastically. For instance, germanium-antimony-tellurium (GST) chalcogenide glass has recently been used to demonstrate all-optical, non-volatile GSPM switch [174]. However, GST needs very high temperature to return back to amorphous state, which may limit the reversibility of GST GSPMs. As such, vanadium dioxide (VO₂), a classical transition metal oxide that behaves as an insulator and metal at temperature below and above the insulator-to-metal transition temperature (∼340 K in bulk crystals), respectively, is a good choice to facilitate reconfigurable GSPMs [166, 177, 178]. Very recently, Werner’s group demonstrated a hybrid metasurface platform that achieves electrically triggered multifunctional control in the mid-infrared region by integrating a VO₂ film of nanoscale thickness into an optical metasurface absorber, where the active VO₂ layer is enabled as a function of the electrical current flowing through the nanoengineered metal layer supporting the optical resonances and the electrical control simultaneously (Figure 11C) [166]. When the electrical current (I) increases to 1.20 A, absolute tunings of 80% and 75%, which correspond to 75-fold and 5-fold relative modulations relative to the static reflectance when I=0 A, are observed at wavelengths of 3.05 and 3.85 µm, respectively. In addition to the drastic tuning, a continuous spectrum tuning is observed for in-between current values before saturation. The proposed electrically active GSPM was further used to tune the spatial dependence of infrared signals, verifying a more sophisticated spatial modulation scheme in the infrared range.

In addition to graphene and phase-change materials, a variety of other functional materials has been successfully used to realize active GSPMs, for instance, transparent conducting oxides [167, 180, 181]. Among the transparent conducting oxide materials, indium tin oxide (ITO) has been widely used as significant voltage-tunable changes in the refractive index can be achieved through accumulation of carriers through the electric field effect when applying an electrical bias between metal and ITO. In particular, the electric field can change the real part of the permittivity of ITO in the accumulation layer from positive to negative, lending to large electric field enhancement in the epsilon-near-zero region. Here we would like to highlight a gate-tunable GSPM that enables dynamic electrical control of the phase and amplitude of the reflected plane wave [167]. As shown in Figure 11D, the metasurface is composed of a continuous gold film and a thin ITO film followed by a thin alumina film on which the connected gold stripe nanoantenna array is patterned. Because of the field-effect modulation of the complex refractive index of ITO layers incorporated into metasurface antenna elements, a phase shift of 180° and ∼30% change in the reflectance was experimentally measured by applying a 2.5-V gate bias. Furthermore, by electrically controlling subgroups of metasurface elements properly, electrical switching of ± first order diffracted beams has been achieved. When the applied voltage is increased above ∼1.5 V, the ±1 order diffracted beams appear and become dominant while the zero-order diffracted beam intensity reduces, in good agreement with the simulation results (right panel of Figure 11D).

As a final comment, it should be noted that while here we are mainly focused on the reconfigurable GSPMs in high frequencies, similar reconfigurability can be
Figure 12: Other applications.
(A) Left upper panel: unit cell of the GSPM. Left lower panel: the average normalized reflectivity along the $x$-coordinate for experimental (solid line) and expected (dashed line) responses. Middle panel: SEM image of integrator GSPM. Right panel: Measured and simulated reflected far-field intensity for integrator metasurface. Reprinted from Ref. [187]. (B) Left panel: a 3D illustration of GSPM cloak for visible light. Right panel: schematic of MIM nanoantenna (left top), SEM image of the fabricated metasurface (right top), and the obtained reflection images when cloak is on (left lower) and off (right lower). Reprinted from Ref. [189]. (C) Left panel: schematic illustration of a single GSPM for multichannel polarization-controllable OAM generation. Right panel: schematic of off-axis multi-OAM generation under illumination of RCP light. Reprinted from Ref. [190]. (D) Left panel: schematic of the metal/MQWs hybrid metasurface design exhibiting giant nonlinear response. Right panel: nonlinear response from the hybrid metasurface. Reproduced from Ref. [191].
achieved by varying the voltages applied on the varactor diodes incorporated in the GSP meta-atoms at low frequencies [182–184].

### 3.9 Other applications

As a rapidly developing research area, GSPMs are expanding very fast every day, and novel functionalities continue to be explored benefiting from their unprecedented control of EM waves and recent advances in nanofabrication technologies. Recently, compact analog computing based on metasurfaces has emerged as a promising candidate for real-time and parallel continuous data processing [185–188]. Motivated by the theoretical work [185], the first metasurface that allows for mathematical operations on visible light was demonstrated using GSPM working in a reflection setup (left panel of Figure 12A) [187]. Proof-of-concept experiments have been conducted to realize a high-quality poor-man’s integrator operating at the wavelength of 800 nm, as shown in the right panel of Figure 12A.

The cloaking device that can cause objects to be partially or wholly invisible to the incoming waves has attracted a great deal of interest, becoming one of the most important applications of GSPMs. It was numerically proved that a carpet cloak can be achieved to mimic the reflection from a planar ground by compensating the local phase with gradient metasurfaces, thereby hiding arbitrary objects in a predesigned region [192, 193]. Lately, an ultrathin invisibility skin cloak for visible light was experimentally demonstrated by MIM resonators with spatially distributed phase shifts (Figure 12B) [189]. With the cloak off, a strong contrast between the cloaked region and the surrounding reflective surface revealed the object, but with the cloak on, the contrast became indistinguishable, rendering the object completely invisible. In contrast to bulky cloaks realized via 3D metamaterials, the GSPM cloak was only 80 nm thick. Because of the scalable feature of the metasurface, the GSPM-based cloak can be extended to microwave for full polarization [194, 195].

Metasurface-based vortex-beam generators have recently gained tremendous interest because of potential applications in optical trapping, high-resolution microscopy, and quantum communication technology [19, 196]. A variety of GSPMs has been demonstrated to generate vortex-beams with arbitrary orbital angular momentum (OAM) values [134, 153, 190, 197]. Very recently, Chen’s group proposed and experimentally demonstrated an efficient approach to generate multiple OAM states with different topological charges in four separate channels and arbitrarily control their superpositions using PB phase GSPM (Figure 12C) [190]. By changing the helicity of incident light from RCP to LCP, the positions of all OAM beams are changed to their centrosymmetric positions with respect to the axis of incident light. Meanwhile, the signs of all topological charges are changed from plus to minus.

In addition to unprecedented applications relying on linear light-matter interaction, metasurfaces offer great capability of designing nonlinear photonic devices with remarkable performance, which is ascribed to the unique ability to localize EM fields at nanoscale volumes, together with significantly relaxed phase matching condition due to subwavelength thickness [25, 26]. In this sense, the combination of strong field enhancement within the gap area and intrinsic nonlinearities of metals make GSPM an excellent platform to realize effective nonlinear optical processes [198–202]. To further enhance the nonlinear response, metal/multiple-quantum-well (MQW) hybrid metasurfaces were experimentally demonstrated by coupling EM mode in GSPM with intersubband transitions of MQW semiconductor heterostructure [191, 203, 204]. The GSPMs not only enhance both the fundamental and second harmonic frequencies but also tailor the near-field polarizations, converting giant nonlinear susceptibility of MQW heterostructures and thus resulting in giant nonlinearity (Figure 12D) [191].

### 4 Conclusions and outlook

The research filed for GSPMs has been rapidly growing as a platform for novel EM properties and functionalities for the reflected light, because of its simplified realization that only requires a single step of lithography. In this paper, we have briefly reviewed the fundamental physics and the latest developments of representative applications in the research area of GSPMs.

Owing to the unprecedented control over reflected light with surface-confined planar components, GSPMs are expected to have an important impact in other promising areas converting both fundamental research and real applications, beyond what we have discussed in this review article. For example, GSPMs provide an ideal platform to create multifunctional metadevices, which efficiently integrate multiple functionalities into a single ultrathin device with compact footprint [153, 205]. Additionally, the so-called quantum metasurfaces, which combine existing quantum systems with metasurfaces, is an emerging area of research [29, 206, 207].
Acknowledgments: This work was funded by the European Research Council (the PLAQNAP project, grant 341054) and the University of Southern Denmark (SDU2020 funding).

References


