The complex band structure of surface plasmons on a square metal hole array

THESIS

submitted in partial fulfillment of the
requirements for the degree of

BACHELOR OF SCIENCE
in
PHYSICS

Author :	A.N. van Delft
Student ID :	1121367
Supervisor :	Dr. M.P. van Exter
2^{nd} corrector :	Prof. Dr. M.A.G.J. Orrit

Leiden, The Netherlands, July 22, 2014

The complex band structure of surface plasmons on a square metal hole array

A.N. van Delft

Huygens-Kamerlingh Onnes Laboratory, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

July 22, 2014

Abstract

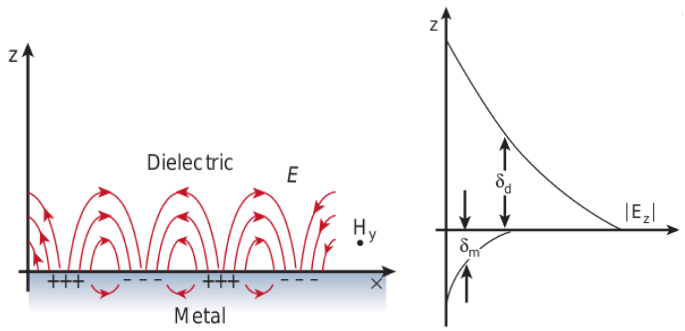
We study the complex band structure of surface plasmons traveling on a gold-semiconductor interface perforated with a two-dimensional square array of holes. To excite the surface plasmons, we use a pump spot with a top-hat profile. The measured spectra contain Lorentzian resonances, which are fitted and compared to a mathematical model. This model explained most of our observations, but still leaves room for improvement.

Contents

Chapter 1

Introduction

Surface plasmons (SPs) are electromagnetic waves combined with a surface charge distribution, traveling on the interface between a metal and a dielectric material, see figure 1.1a. The strength of their electric and magnetic fields decay exponentially over distance, with a decay length characterized by the type of metal and dielectric medium (figure 1.1b).



(a) ***(b)***

Figure 1.1: ***(a)*** *The combined electromagnetic wave and surface charge character of SPs on the interface of a metal and dielectric medium. At the interface, the electric field E is pointing outwards in the z direction and the magnetic field H in the y direction.* ***(b)*** *The strength of the electric (and magnetic) field decreases exponentially over distance, characterized by the decay length determined by the media involved. In general, the decay length δ_d of the dielectric medium is much larger (of the order of half the wavelength of light involved) than the decay length δ_m of the metal (which is determined by the skin depth). Both pictures are taken from ref. [1].*

We study the complex band structure SPs traveling on a a gold-semiconductor

interface perforated with a two-dimensional square array of holes. The understanding of so called *plasmonic crystals* helps us understand and control the light flow through media by, for example, forming mirrors or waveguides[2]. The possibility to add optical gain has also led to lasing action at nanoscales[3, 4], much smaller than possible with current lasers.

The structure of this thesis is as follows. In chapter 2 we will describe the experimental setup and the top-hat profile of the pump bundle. In chapter 3, we will develop some theory and briefly summarize the *coupled-mode model*. In chapter 4 we will show our main measurements and fit Lorentzian type peaks to them. We will compare these fits with the coupled-mode model. In chapter 5 we will show some notable things found in other arrays, including some unexpected asymmetries. And finally, we will have a concluding discussion in Chapter 6.

There are also two appendices: appendix A, which contains a detailed description of the Python code developed to fit the Lorentzian resonances, and appendix B, which contains our article currently under review entitled *Loss and scattering of surface plasmons on optically-pumped hole arrays*. The article contains a more extended description the main results from chapter 4 and the coupled-mode model, as briefly mentioned in chapter 3.

Chapter 2

Experimental setup

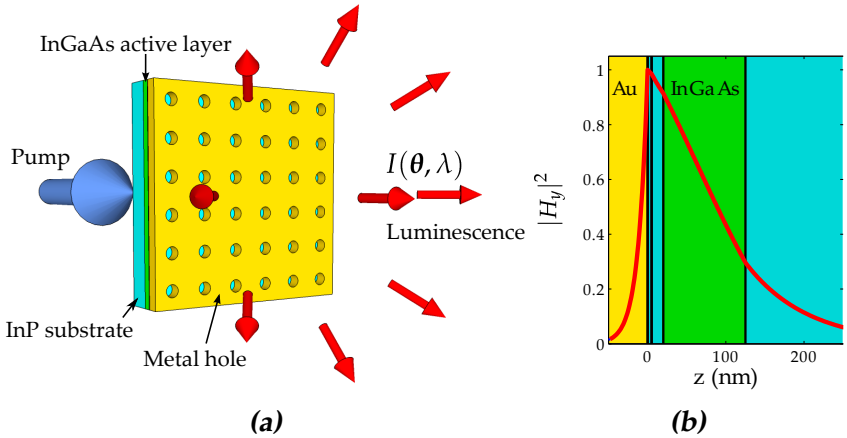
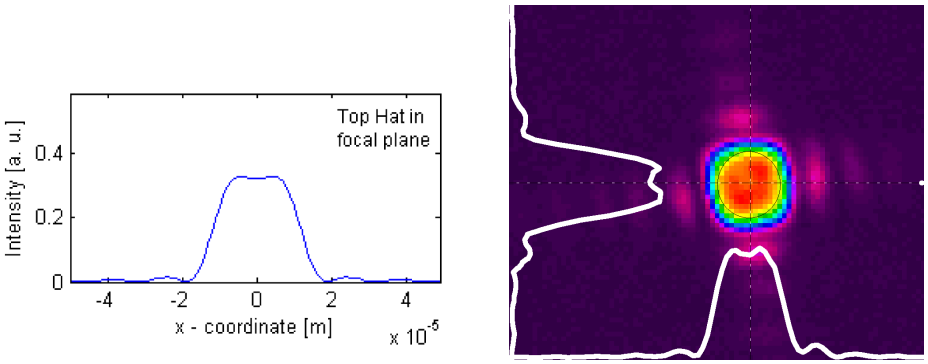


Figure 2.1: ***(a)*** *Geometry of our setup. The pump bundle creates electron hole pairs in the InGaAs gain layer, which recombine radiatively and couple into SPs. The SPs will radiate outwards to the external photon field through the dipoles they induce at the holes. We measure the emitted spectra as a function of the angle* $\boldsymbol{\theta} = (\theta_x, \theta_y)$. ***(b)*** *Schematic view of the layered structure of the arrays. The red line shows the strength of the magnetic field* H_y *of the SPs, which decreases exponentially over distance from the interface. The SPs travel on the Au - InGaAs interface. The Au - air interface is protected with a Chromium layer, to prevent SPs from traveling here. The blue layers are the InP substrate and a thin InP layer to prevent quencing of the excited carriers in the InGaAs to the gold. Both pictures are taken from ref. [4].*

2.1 Experimental geometry

Figure 2.1 shows the geometry of our sample. The sample is 50 × 50 µm and the SPs travel at the gold - semiconductor interface. The semiconductor acts as a gain layer, to amplify the traveling SPs. The gold layer has a square lattice of subwavelength holes, with lattice spacing varying from 450 to 480 nm, in which magnetic dipoles are induced at the holes when SP waves travel by. These dipoles in turn radiate outwards (to the air side as well as the InGaAs). We measure the spectra of the emitted light on the air side as a function of the outcoming angle $\boldsymbol{\theta} = (\theta_x, \theta_y)$. The light is captured with a 20× microscope objective with NA = 0.4.

2.2 Pump bundle (top-hat profile)



(a) ***(b)***

Figure 2.2: ***(a)*** *Cross-section of top-hat profile as specified by TOPAG. The height of the side peaks is about 6% of the maximum and the central dip is 2%. The figure is taken from the TOPAG manual.* ***(b)*** *2D profile of the top-hat, as created using the setup of figure 2.3. Here the height of the side peaks is 13% of the maximum and the central dip is 6%, so it's a bit less flat than the manual states. This is due to the beam size on the hologram, which is apparently slightly bigger.*

We used a 1064 nm laser as a pump to excite the gain layer. To distribute the light more equally over the whole sample, we used a commercial hologram (the TOPAG GTH-4-2.2) which turned our laser with a Gaussian profile into a square top-hat profile, as depicted in figure 2.2. Figure 2.3 shows the setup used to get a good top-hat focus on our sample. The focus had a size of around 50 × 50 µm, so it's approximately as big as a sample.

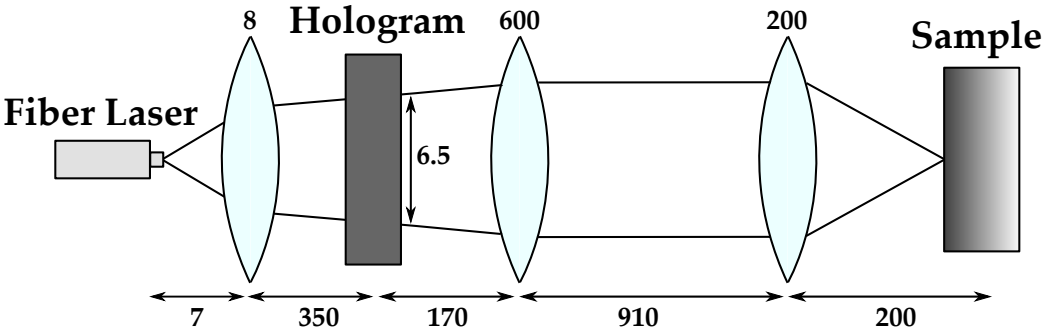
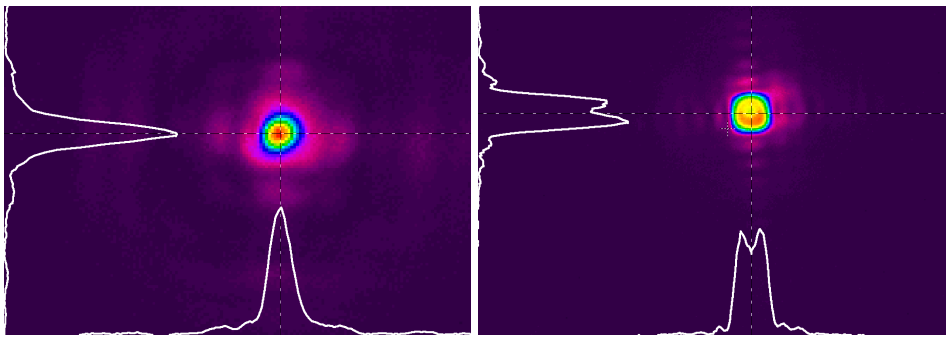


Figure 2.3: *Schematic view of the setup used to convert a Gaussian laser beam into a top-hat profile as in figure 2.2. Because the width of the bundle had to be close to 6.5 mm, we placed a lens after the laser to have some more space. After going through the hologram, the beam is focussed on the sample by the last two lenses. All lengths are in mm.*

2.2.1 Alignment

We encountered two challenges in aligning the setup in figure 2.3. The first is making sure the beam on the hologram is indeed 6.5 mm big, which had to be really precisely tuned. If the beam was too big, the center jump in the flat-top profile would be too low, and if the beam was too small, the outgoing beam would not be effected enough and it would still have a Gaussian profile (see figure 2.4). A square engraved on the hologram glass helped getting the right size of the beam.



(a) ***(b)***

Figure 2.4: *2D profiles of the focus of beams that were* ***(a)*** *too small or* ***(b)*** *too big on the hologram. If the beam is too small, the hologram doesn't influence the beam very much and it would still approximately be a Gaussian profile. If it is too big, the central dip of the flat-top would be too low (here it is 20%).*

The second challenge is the alignment of the last 200 mm lens, to illumi-

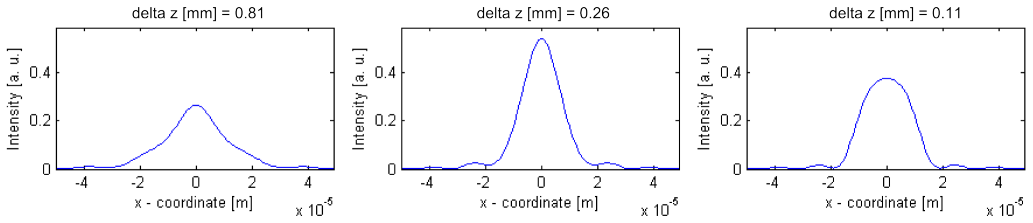


Figure 2.5: *Cross sections of the beam profile slightly out of focus, the distance from focus decreases from left to right. We see a peak coming up which turns into a jump when in focus (figure 2.2a). The profile is approximately symmetric with respect to the focal plane, so on both sides of the focus this pattern can be seen. Again, the figures are taken from the hologram manual.*

nate the sample in the focal plane. Out of focus (on both sides), the central dip of the top-hat would change in a peak, see figure 2.5. This profile is approximately symmetric with respect to the focal plane. So to make sure we were in focus, we had to be exactly in between the peaks arising at both sides of the focal plane.

Chapter 3

Theory

Before describing our measurements, we want to take a brief theoretical detour in order to understand SPs and their interaction with the metal hole array. In this chapter, we will briefly discuss the *coupled-mode model*, which is explained in more detail in appendix B.

3.1 Theory of SPs on a flat interface

SPs are electronic and optical solutions to the macroscopic Maxwell equations. Let ε_d and ε_m be the frequency-dependent dielectric functions of the dielectric medium and metal respectively, which in general are complex functions. From the continuity of the boundary conditions at the interface, we can derive the dispersion relation

$$\tilde{k}_{SP} = \frac{\omega}{c}\sqrt{\frac{\varepsilon_d \varepsilon_m}{\varepsilon_d + \varepsilon_m}} \tag{3.1}$$

for SPs on the interface[1]. Here, the real part of $\tilde{k}_{SP}$ denotes the magnitude of the wave vector $\mathbf{k}_{SP}$ and the imaginary part the *loss* or *penetration depth* on the interface.

3.2 The coupled-mode model

In ref. [5], a simple mathematical model for the SPs that exist on the metal-dielectric interface of a square hole array was created. This model is extended in appendix B. We will give a brief summary of this model and use it to formulate our expectations for our measurements.

3.2.1 Brief summary of the model

The model approximates the SP-field on the interface by its first order Fourier terms which correspond to traveling waves in the $\pm x$ and $\pm y$ direction. It uses the dispersion relation (3.1) of SPs on a flat interface and considers coupling to emitted photons with (in-plane) parallel momentum $\mathbf{k}_{\parallel} \approx \omega/c(\sin\theta_x \mathbf{e}_x + \sin\theta_y \mathbf{e}_y)$ originating from SPs with momentum $\mathbf{k}_{SP} = \mathbf{k}_{\parallel} + \mathbf{G}_i$[6], where $\mathbf{G}_i$ is a lattice vector of the hole array (again only first order Fourier terms with $|\mathbf{G}_i| = 2\pi/a_0$ in the $\pm x$ and $\pm y$ direction are used). At any point in time we can write the SP field $|E\rangle$ as a superposition of the four traveling waves and consider SP-to-SP scattering between those four modes by a scattering matrix H. We then solve the differential matrix equation

$$\frac{d}{dt}|E\rangle = -iH|E\rangle \tag{3.2}$$

to describe the interaction between different modes. Appendix B describes what H looks like and how to solve this differential equation.

3.2.2 Expectations for $\theta_x = 0$

For the restriction $\theta_x = 0$, which we use in most of our results, we expect four different θ_y-dependent eigenvalues, called *bands*. The inclusion of the gain of the semiconductor, ohmic losses and losses to the external photon field in our model results in *Lorentzian resonances* around these eigenfrequencies. Three of those bands are expected to radiate in the *P-polarization*, with the electric field in the plane of incidence (the yz-plane), and one in the *S-polarization*, perpendicular to the plane of incidence. Furthermore, the S-polarized mode and one of the three P-polarized modes are expected to be *bright*, which means they radiate near $\theta_y = 0$, and the other two P-polarized modes are expected to be *dark*, which mean they do not radiate here. The distinction between bright and dark modes is related to the *cosine* and *sine* nature of the standing waves in the $\pm x$ and $\pm y$ direction.

In the next chapter, we will show simulations of the Lorentzian resonances based on the coupled-mode model. We will compare these simulations to our data. The model includes a couple of parameters: the right-angle scattering rate κ, the backwards scattering γ, the gain of the InGaAs layer and various loss channels. We use the same scattering rates as in ref. [5] and the gain and loss channels are deduced from the spectral widths of the Lorentzian fits to our data (see appendix B).

Chapter 4

Main results

We studied the emitted optical spectra of arrays with different lattice spacing a_0, varying from $a_0 = 450\,\text{nm}$ to $480\,\text{nm}$. In this chapter we will focus on one array in particular, with $a_0 = 470\,\text{nm}$. This array is from a sample created by Frerik van Beijnum, which has also been used in ref. [4, 5]. The measurements done on this array are thoroughly discussed our article in appendix B. We measured emitted light spectra at $\theta_x = 0$ and scanned over θ_y (figure 4.1a) from -0.3 to 0.3 radians (rad). Other measurements and typical observations are shown in the next chapter.

4.1 Raw measurements

We measured the emitted light in the P- and S-polarization. The angle dependent spectra can be displayed as a two dimensional intensity spectrum $I(\theta_y, \lambda)$ in both polarizations, shown in figure 4.1b and 4.1c.

4.2 Lorentzian fits

The developed theory described in the previous chapter predicts Lorentzian type resonances in ω (the frequency of the emitted light). The spectra are measured as a function of the wavelength $\lambda \propto \omega^{-1}$, but one can show that for small widths ($\Delta\lambda \ll \lambda_0$), the peaks in λ are also approximately Lorentzian. Lorentzian peaks are of the form

$$I(\lambda) = \frac{A}{\pi} \frac{\Delta\lambda}{(\lambda - \lambda_0)^2 + \Delta\lambda^2} \tag{4.1}$$

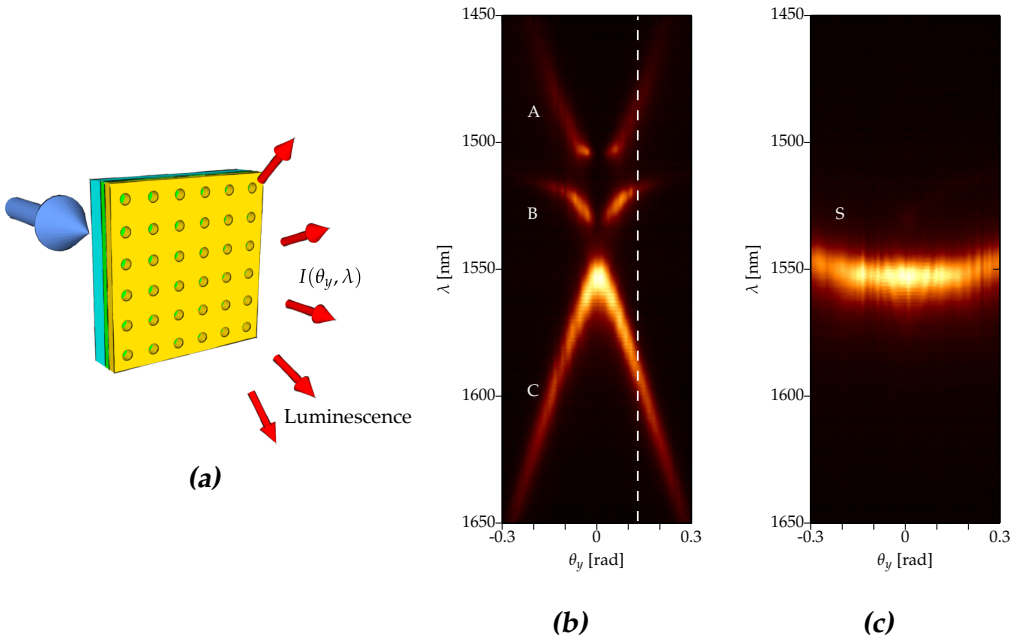


Figure 4.1: ***(a)*** *The experimental geometry, where the reduced set of arrows emphasize the fact that we are looking in the $\theta_x = 0$ region and scanning over θ_y.* ***(b)*** *,* ***(c)*** *False color plots of the emitted P- and S-polarized spectra respectively as a function of the emission angle θ_y. The dashed line in the P-polarized plot indicates the location of the cross-section showed in figure 4.2. In both polarizations we observe a total of four bands, labeled A, B, C (P-pol) and S (S-pol), as indicated in the figure.*

and can thus be characterized by three parameters: the central wavelength λ_0, the spectral width (HWHM) $\Delta\lambda$ and the integrated intensitiy (area under the curve) $A = \pi I_{max} \Delta\lambda$, where $I_{max} = I(\lambda_0)$. For every cross section at a given θ_y we fit an incoherent sum of Lorentzians, as depicted in figure 4.2. Each fit yield three fit parameters per Lorentzian peak. For more details on the fit procedure and the Python code, head to appendix A.

We then combine all the fit results obtained at different values of θ_y. For each of the four bands depicted in figure 4.1b and 4.1c, this results in the angle dependency of the central wavelength, spectral width and integrated intensity of the fitted Lorentzian peaks. In addition to these upcoming *parameter plots*, we will also show the simulations of these angle dependenent parameters, as mentioned in chapter 3.

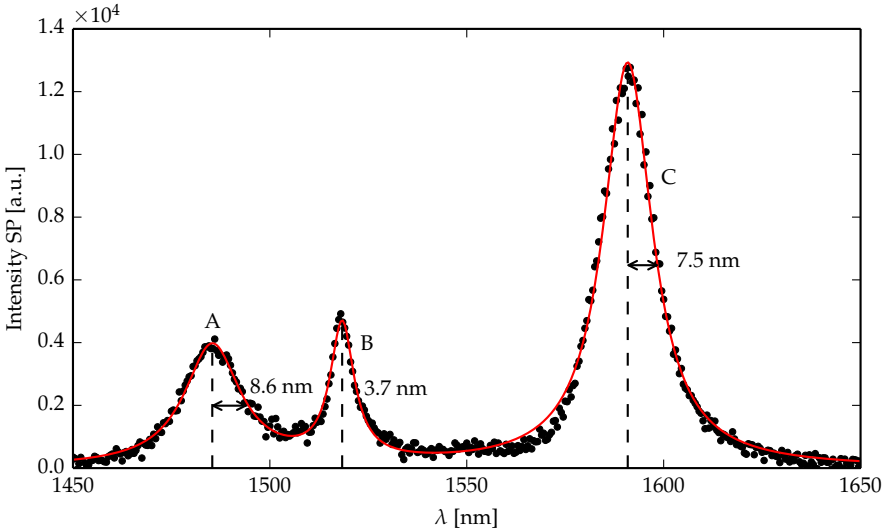


Figure 4.2: *Cross section of the P-polarized spectrum at $\theta_y = 0.13$ rad, as indicated by the dashed line in fig 4.1b. We fitted an incoherent sum of three Lorentzian type resonances through the data as indicated by the red line. These fits returned for each peak the central wavelength, spectral width and area under the curve.*

4.2.1 Central wavelengths

Figure 4.3 shows the fitted central wavelengths of all four bands, together with the simulations. By comparing the data to the simulations, we see striking similarities. The biggest discrepancy are the kinks in the A and S band at $\theta = \pm 0.06$ rad and $\theta = \pm 0.17$ rad, respectively, as indicated by the arrows. Incidentally, these kinks lie exactly on a light line of the semiconductor. This gives us a possible explanation, which we haven't incorporated in our model. We will give this explanation after showing the angle dependency of the spectral widths, which also show unexpected behavior at those points.

4.2.2 Spectral widths

Figure 4.4 shows our fitted and simulated spectral widths. The unit chosen is the dimensionless unit $\Delta\omega/\omega_0$, which for small widths ($\Delta\lambda \ll \lambda_0$) is equal to $\Delta\lambda/\lambda_0$. We observe strange jumps in the S band at $\theta_y = \pm 0.17$ rad, exactly the same value as the location of the kinks observed in figure 4.3. The explanation we have for this is the occurence of a second order diffraction, which appears beyond the light lines of the semiconducter as

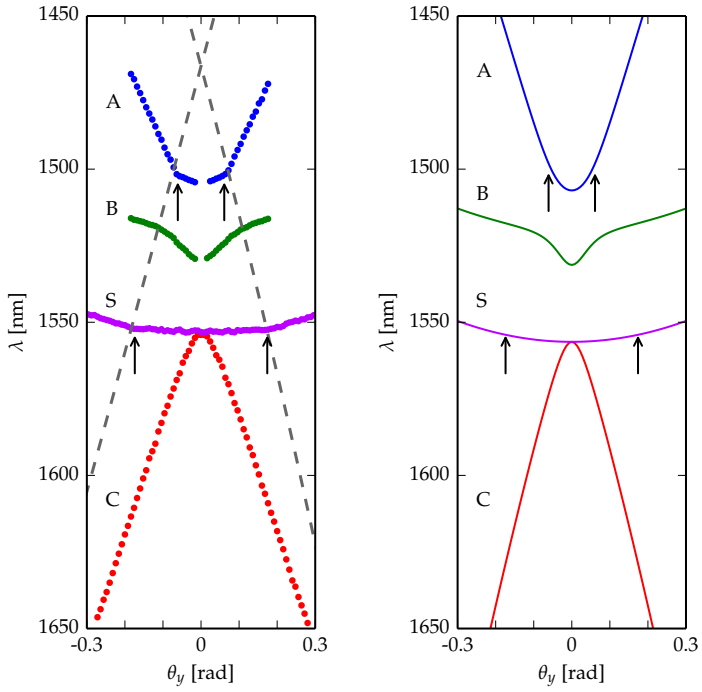


Figure 4.3: ***(a)*** *The fitted central wavelengths of the Lorentzian resonances as depicted in figure 4.1. The arrows indicate some strange kinks in the A and S band. The dashed lines are light lines of the semiconductor mode. All the kinks lie on these light lines.* ***(b)*** *Simulations of the central wavelength, based on the model explained in chapter 3. The arrows indicate the places were the kinks should be, but they do not appear.*

indicated in figure 4.3 (see appendix B for more details). This extra loss-channel results in a sudden jump in the spectral width, which is exactly what we see. This phenomenon is called a *Rayleigh anomaly*[7]. In the A band this jump is not very clear, because the losses are increasing quickly. The problem here is that we are near the Fermi energy, where the gain of the semiconductor becomes negative, and thus becomes a loss, which makes it hard to see a clear jump. In chapter 5 we mention another array that shows a much clearer jump in the A band. The jumps in the spectral widths in turn affect the central wavelengths because they are related via the Kramers-Kronig relations to the spectral widths[6, 8].

The simulations of the spectral widths do not match the measurements very well. Near the center we see similarities in the shape of the A and B

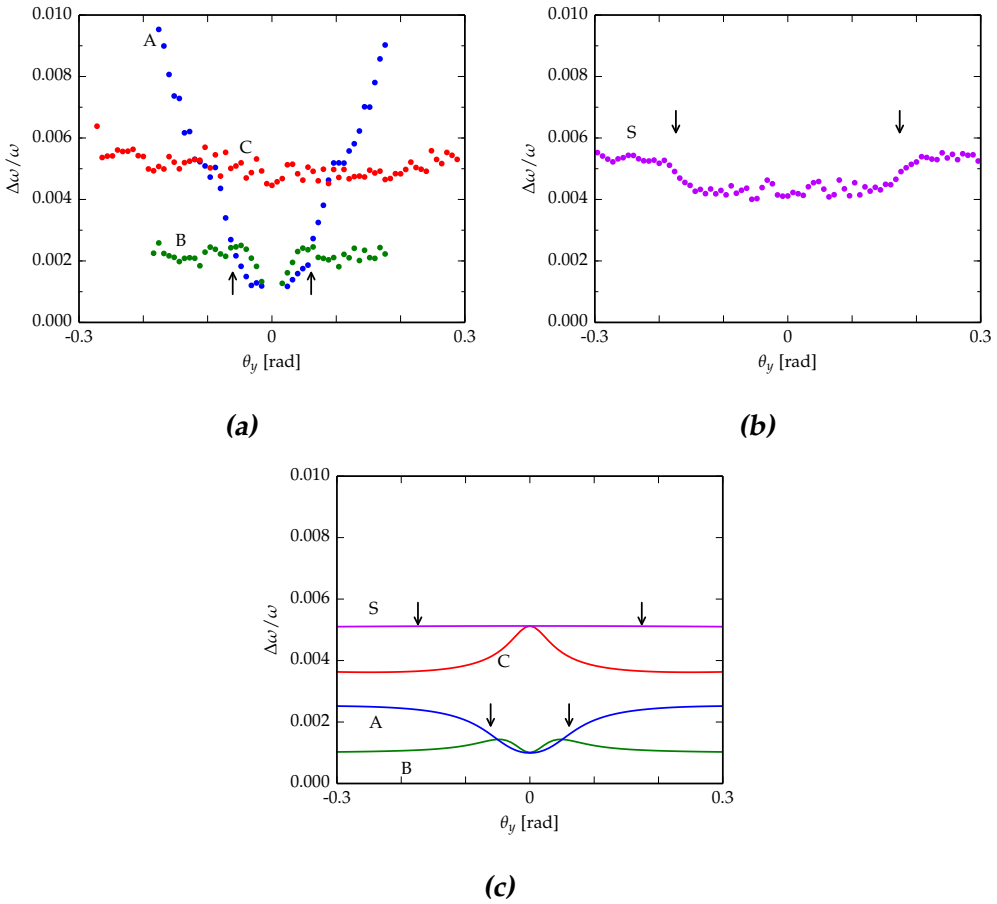


Figure 4.4: ***(a)*** *,* ***(b)*** *Parameter plot of the spectral widths as fitted from the Lorentzian resonances in figure 4.1, in the P- and S- polarization, respectively. The arrows mark the locations of the kinks in figure 4.3a.* ***(c)*** *Simulations on the spectral widths, based on the model explained in chapter 3, again with arrows on the locations of the kinks in figure 4.3a.*

band, but off center this is less accurate. Furthermore, our measurements do not show the expected peak in the center of the C band as predicted by our simulations.

4.2.3 Integrated intensities (areas)

Figure 4.5 shows the integrated intensities A. Our model in chapter 3 predicts two bright and two dark modes at the center, associated with the sine or cosine type standing waves in the x and y direction. We indeed observe this in figure 4.1, and figure 4.5 quantitatively confirms this. Also,

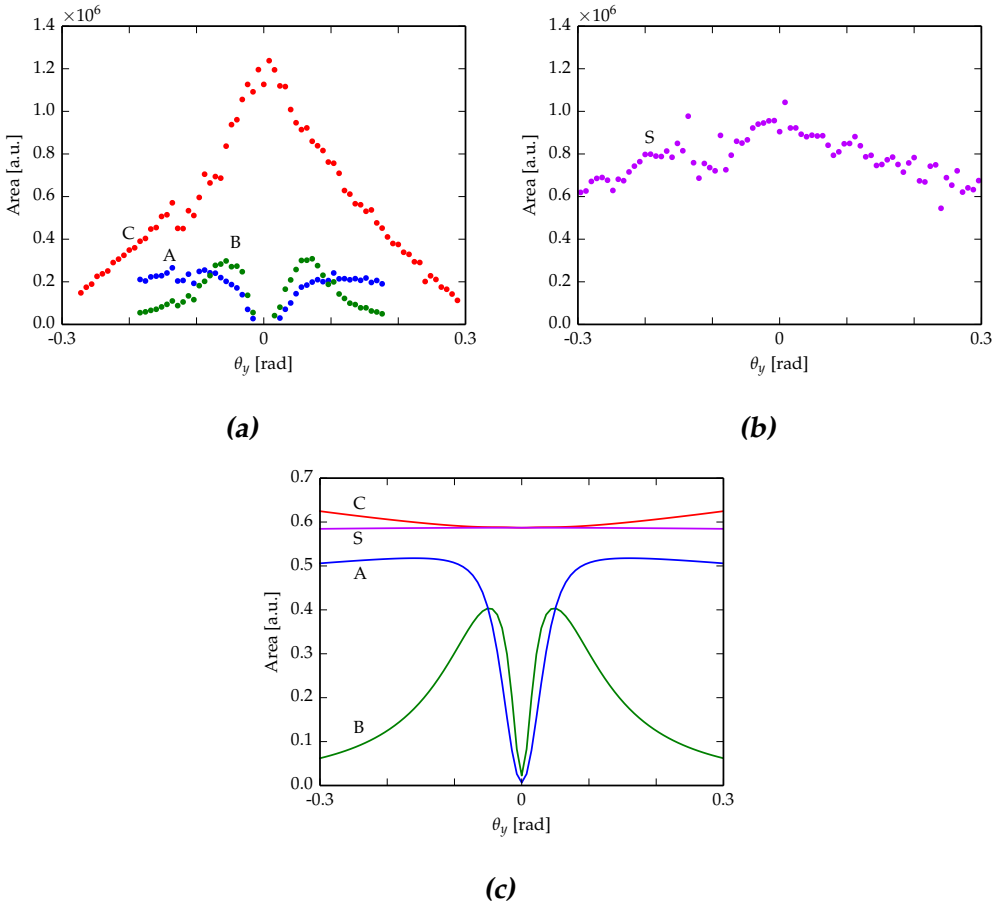


Figure 4.5: ***(a)*** *,* ***(b)*** *Parameter plots of the integrated intensities as fitted from the Lorentzian resonances in figure 4.1.* ***(c)*** *Simulations of the integrated intensities, based on the model explained in chapter 3.*

the shapes of the B and C band approximately match those of the simulations, although their relative intensities are off. Other than this, there is little agreement with our simulations.

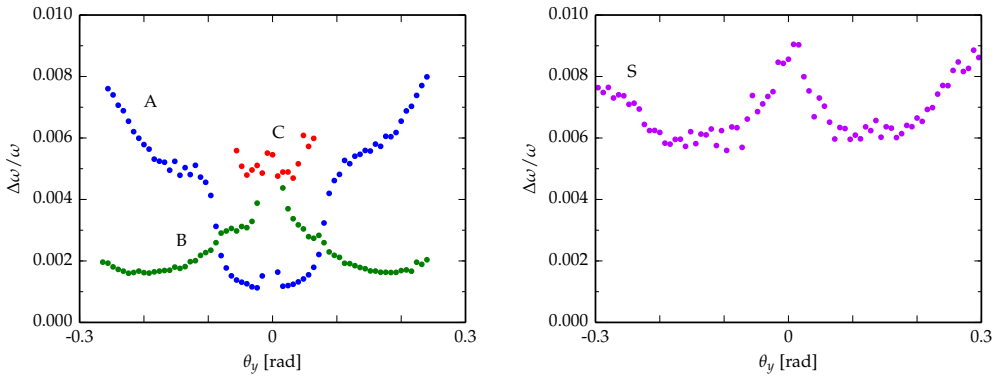
Chapter 5

Results for other arrays

In this chapter we will discuss results obtained for other arrays. In general, all the features presented in the preceding chapter, like the jumps in the spectral widths and kinks in the central wavelengths, were observed in all our measurements. There are a few special cases that we want to highlight. We will show an array that shows a much clearer jump in the spectral widths of the A band (again at $\theta_x = 0$) than in figure 4.4, but also shows a strange and unexpected bump near normal incidence ($\theta_x = \theta_y = 0$) in the spectral widths of the S band. Also, we will highlight emission spectra of arrays that violate some expected symmetries. All arrays analyzed in this chapter are from a different sample than the array of chapter 4. This sample was created by Vasco Tenner (labeled with RA2083-1-b).

5.1 Features in the spectral widths

In subsection 4.2.2 we discussed the expected jump in the A band which was hard to see, due to the decreasing gain. We measured the intensity spectra $I(\theta_y, \lambda)$ on an array from Vasco's sample with lattice spacing $a_0 = 480\,\text{nm}$, in which this jump *can* be seen clearly before the widths are increasing. Figure 5.1 shows the parameter plots of the spectral widths in the S- and P-polarization. The S band again showed a (slightly smaller) jump. But in contrast to our earlier measurements in chapter 4, we see a strange bump near normal incidence in the S band.



(a) ***(b)***

Figure 5.1: ***(a)*** *Fitted spectral widths of the Lorentzian resonances emitted in the P-polarization by an array with $a_0 = 480$ nm from Vasco's sample. The expected jump in the A band can be seen much better than in the array by Frerik used in chapter 4, because it is much further from the Fermi energy.* ***(b)*** *Fitted spectral widths of the Lorentzian resonances emitted in the S-polarization by the same array. This also shows the expected jump. However, near normal incidence, a strange bump can also be seen. This feature has not been observed before and is not yet understood.*

5.1.1 More bumps

The measurements on this array belonged to a series of measurements on different arrays from Vasco's sample. In all the measurements the bump in the spectral widths of the S band are observed, whereas earlier measurements, which were taken from Frerik's sample, do not show these bumps. This suggests that the bumps are some sort of artefact of the sample itself which perhaps occurred at the creation of the sample.

5.2 Asymmetric arrays

Due to the geometry of our array (figure 2.1), we are dealing with mirror and rotational symmetries in our measurements. So far these symmetries were conserved (for example the mirror symmetry $\theta_y \to -\theta_y$ is conserved in all our previous measurements). However, measurements on some arrays of Vasco's sample showed the breaking of some expected symmetries. We found unexpected asymmetries in the P- and S-polarized light at normal incidence and in the measured spectra of some 2D scans (in which we scan over both θ_x and θ_y).

5.2.1 Asymmetry in P- and S-polarization at normal incidence

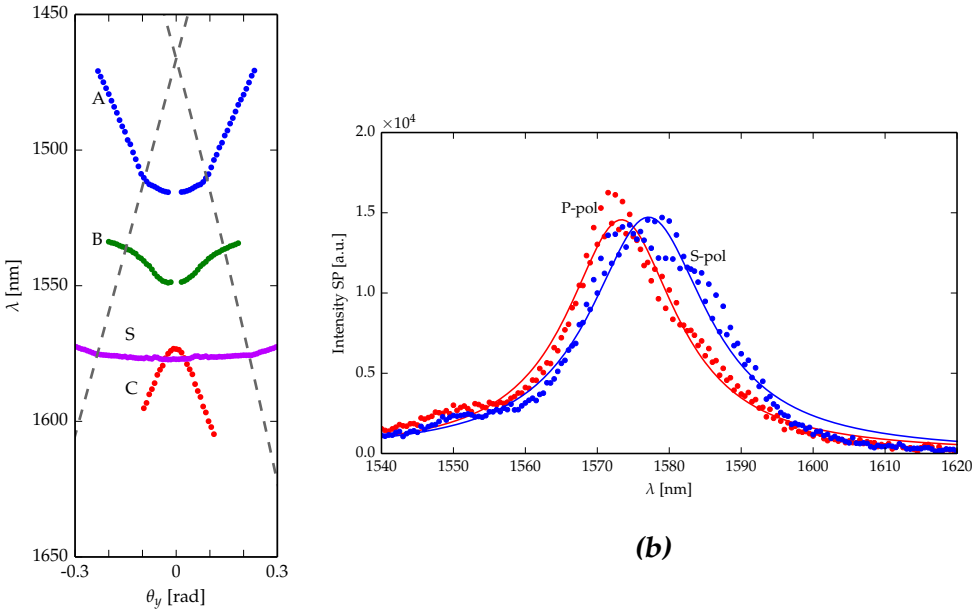


Figure 5.2: ***(a)*** *Central wavelength parameter plot of an array with $a_0 = 470$ nm, from Vasco's sample. We observe an unexpected asymmetry at $\theta_y = 0$* ***(b)*** *Cross section of both polarizations at $\theta_y = 0$, which confirms the asymmetry observed in the central wavelength plot. The red data is the P-polarized light and the blue the S-polarized light, including fits matching their color.*

If we look at light emitted at normal incidence, we expect the P- and S-polarized light to be the same*, due to the $\pi/2$ rotational symmetry of our geometry. However, measurements on some arrays did show an asymmetry between the two polarizations, see figure 5.2. There were some other arrays where this phenomenon is also observed, again only on Vasco's sample. On one of those measurements we did a 2D scan to look for other notable features, in which we encountered asymmetries in the xz- and yz-plane.

*Strictly speaking the P- and S-polarization can not be defined at normal incidence, because there is no plane of incidence. But because we are scanning over θ_y, we define the yz-plane as our plane of incidence, such that the P- and S-polarization can be clearly defined.

5.2.2 Asymmetric 2D scan

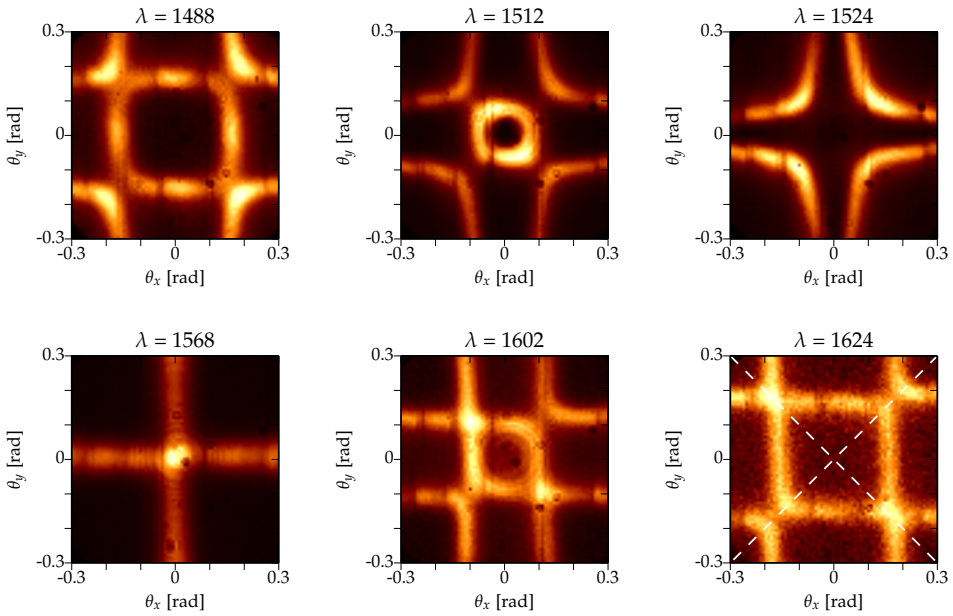


Figure 5.3: *Intersections at different wavelengths of the three-dimensional intensity plot $I(\theta_x, \theta_y, \lambda)$. Every intersection should have the mirror symmetries in the xz- and yz-plane, but we observe a violation in this symmetry. As the dashed lines of the last plot indicates, the mirror symmetries in both diagonal planes are still conserved. The last two plots show a rhomb at the center with a small angle of 86°.*

Again due to the geometry of our arrays, we expect our measurements to have a mirror symmetry in the xz- and yz-plane. So we measured the emission spectra with both θ_x and θ_y varying from -0.3 to 0.3 rad on the same array used in subsection 5.2.1. This generates a three dimensional intensity plot $I(\theta_x, \theta_y, \lambda)$. Figure 5.3 depicts intersections of this intensity plot at different values of λ. At every intersection we expect the xz and yz mirror symmetry, but the intersections show that these are broken. The profile seems a bit squashed to one diagonal and stretched to the other. However, the mirror symmetries in both diagonal planes *are* conserved. This suggests that the array has a rhomb-shaped lattice as depicted in figure 5.4. Earlier 2D scans on Frerik's sample *did* show the expected mirror symmetries in the xz- and yz-plane[4, 5].

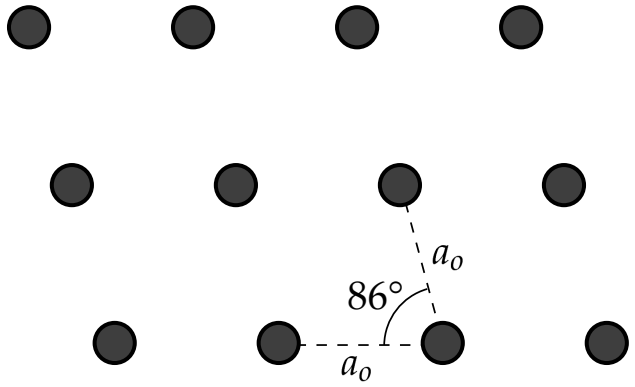


Figure 5.4: *The rhomb-shaped lattice we expect the analysed array to have. A rhomb does not have the xz- and yz-plane mirror symmetries, but only have the mirror symmetries in both diagonal planes, which are still perpendicular to each other. The rhombs observed in the last two plots of figure 5.3 both had a small angle of 86°, which suggests that the angle between the two lattice vectors of our unit cell is also 86°.*

5.3 Connection between bumps and asymmetric samples?

Incidentally, all three arrays that showed bumps in the spectral widths of the S band as in figure 5.1b also had the same symmetry breaking as in 5.2. This suggests a connection between the two phenomena. We could not confirm the mirror symmetry breaking of figure 5.3 on other arrays of Vasco's sample, because we didn't make any 2D scans of this array. We unfortunately do not have more measurements on this sample and also do not have a proper explanation for the connection. Hence, the sample used needs further investigation on these bumps and unexpected asymmetries.

Chapter 6

Concluding discussion

In summary, we performed a detailed analysis on the complex band structure of SPs on square metal hole arrays. We excited the SPs with a top-hat profile pump bundle, using a TOPAG hologram. We fitted incoherent sums of Lorentzian resonances to our data using Python code. Using this, we obtained the angle dependence of the central wavelengths, spectral widths and integrated intensities of the bands. We compared these results to a coupled-mode model.

We analysed two different samples, one created by Frerik van Beijnum and one by Vasco Tenner. The results obtained for Frerik's sample matches our theory reasonably well at some places, but has it's shortcomings at other points. Kinks in the central wavelengths and sudden jumps in the spectral widths were observed, wich we both attribute to a Rayleigh anomaly. Vasco's sample gave similar results, but also showed unexpected bumps in the spectral widths of the S band. some unexpected asymmetries. After investigating the asymmetries, we hypothesize that the hole array is not square but rhomb-like with a small angle of 86°.

We like to conclude with two suggestions for future research. In order to gain a better understanding of the SPs, the coupled model has to be improved to account for the discrepancies observed. Also, Vasco's sample needs further investigation in the bumps and asymmetries, and a possible connection between them. This includes SEM (Scanning Electron Microscope) pictures of the arrays, to investigate their geometry.

References

[1] W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Surface plasmon subwavelength optics*, Nature **424**, 824–830 (2003).

[2] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light (Second Edition)*, Princeton University Press, 2011.

[3] J. Bravo-Abad and F. J. García-Vidal, *Plasmonic lasers: A sense of direction*, Nature Nanotechnology **8**, 479 (2013).

[4] F. van Beijnum, P. J. van Veldhoven, E. J. Geluk, M. J. A. de Dood, G. W. 't Hooft, and M. P. van Exter, *Surface Plasmon Lasing Observed in Metal Hole Arrays*, Physical Review Letters **110**, 206802 (2013).

[5] M. P. van Exter, V. T. Tenner, F. van Beijnum, M. J. A. de Dood, P. J. van Veldhoven, E. J. Geluk, and G. W. 't Hooft, *Surface plasmon dispersion in metal hole array lasers*, Optics Express **21**, 27422 (2013).

[6] C. Kittel, *Introduction to Solid State Physics*, Wiley, 2005.

[7] H. Gao, J. M. McMahon, M. H. Lee, J. Henzie, S. K. Gray, G. C. Schatz, and T. W. Odom, *Rayleigh anomaly-surface plasmon polariton resonances in palladium and gold subwavelength hole arrays*, Optics Express **17**, 2334 (2009).

[8] V. Lucarini, *Kramers-Kronig Relations in Optical Materials Research*, Springer Science & Business Media, 2005.

Appendix A

Python code on the Lorentzian fitting

In this appendix, we will describe the Python code used to fit the Lorentzian peaks to our data as depicted in figure 4.2. We created a class `lorfit` with different members to fit data and to plot data, fits or fitted parameters.

A.1 Code walkthrough

The main function used is `scipy.optimize.curve_fit`, which used a non-linear least squares algorithm to fit a function `f`, with variable parameters, to two-dimensional data. The strategy is to go through every slice of θ_y (associated with a index number `N` in our code) and fit an incoherent sum of Lorentzian peaks as defined in equation (4.1). However, in order to know how many peaks we have to fit, we first have to count them in the data.

A.1.1 Counting number of Lorentzian peaks

The class member `lorfit.findpeaks` counts the number of peaks in a given dataset. This function is defined as:

```
def findpeaks(self,N):
  peaks = np.array(find_peaks_cwt(self.data[N], np.
   array([10])))
  if peaks.shape[0] != 0:
```

```
    peaks = np.sort(peaks[(self.data[N,peaks]>self.
   thresh)])
  return peaks
```

The function returned the coordinates of the maxima of the peaks. We used the function `scipy.signal.find_peaks_cwt`, which locates local maxima in a given dataset. At low Signal-to-Noise Ratio (SNR), many incorrect peaks were found in the noise. For this reason a threshold is inserted (`lorfit.thresh`), below which the found peaks are ignored.

A.1.2 Defining the fit function

We defined sums of Lorentzian peaks as:

```
def lor_0(x):
  return np.zeros(x.shape)

def lor(x, x0,gamma,A):
  return (A/np.pi)*(gamma)/((x - x0)**2 + (gamma)**2)

def lor_2(x, x1,gamma1,A1,x2,gamma2,A2):
  return lor(x, x1,gamma1,A1) + lor(x, x2,gamma2,A2)

def lor_3(x, x1,gamma1,A1,x2,gamma2,A2,x3,gamma3,A3):
  return lor(x, x1,gamma1,A1) + lor(x, x2,gamma2,A2) +
    lor(x, x3,gamma3,A3)

L = [lor_0,lor,lor_2,lor_3]
```

So for a given number of peaks `no_peaks` ≤ 3 is `L[no_peaks]` defined as an incoherent sum of `no_peaks` Lorentzian peaks. A function `L[no_peaks]` thus has $3 \times$ `no_peaks` parameters.

A.1.3 Fitting the Lorentzian peaks

When the number of peaks is determined, the fitting of the Lorentzian resonances can begin. An optional argument of the function `curve_fit` is `p0`, an array with the initial guesses of the parameters of the input function `f`. Because we already located the maxima, we used this information for `p0`.

- For the central wavelengths λ_0 we used the x-coordinate of the found maxima x_{max}.

- $\Delta\lambda$ had a default value of 5 nm but could be altered by the user (via `lorfit.init_width`).

- And finally we estimated the area A to be $\pi y_{max}\Delta\lambda$ with y_{max} the y-coordinate of the found peak. The area is chosen such that the maximum of the initial guess coincides with the found maximum (x_{max}, y_{max}) by the function `lorfit.findpeaks`.

The process explained is done with the code below, the first bit of the code of the class member `lorfit.fit`.

```
def fit(self,N):
  y = self.data[N,:]
  x = self.pixels

  peaks = self.findpeaks(N)
  no_peaks = peaks.shape[0]
  if no_peaks > 3:
    no_peaks = 3 #Max of 3 Lorentzian peaks
  init_param_px = np.zeros(no_peaks*3)
  for i in range(no_peaks):
  init_param_px[3*i] = peaks[i]
    init_param_px[3*i+1] = self.init_width
    init_param_px[3*i+2] = np.pi*y[peaks[i]]*self.
   init_width

  if no_peaks != 0:
    try:
      param_px = curve_fit(L[no_peaks],x,y,p0=
   init_param_px,maxfev=int(1e4))[0]
    except RuntimeError:
      print RuntimeError, 'N =', N
      param_px = np.zeros(3*no_peaks)
    for i in range(3 - no_peaks):
      param_px = np.append(param_px,np.zeros(3))
  else:
    param_px = np.zeros(9)
  param_px = np.sqrt(param_px**2)
```

Figure A.1 graphically summarizes the explained procedure of finding the best fit to the Lorentzian resonances.

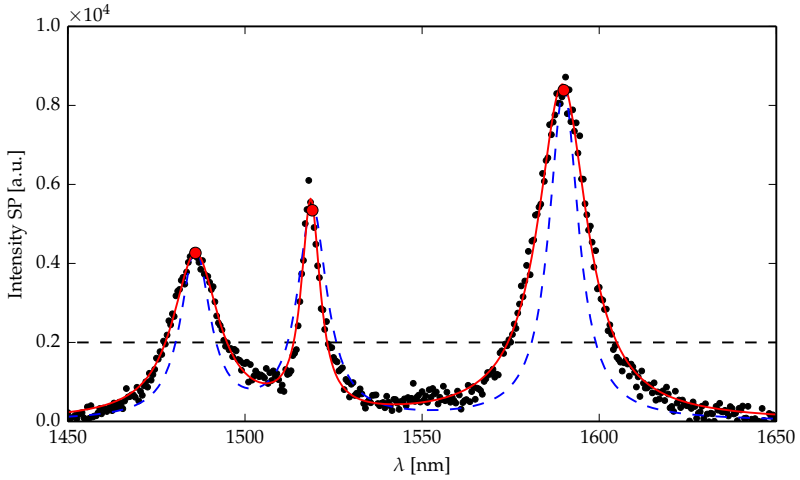


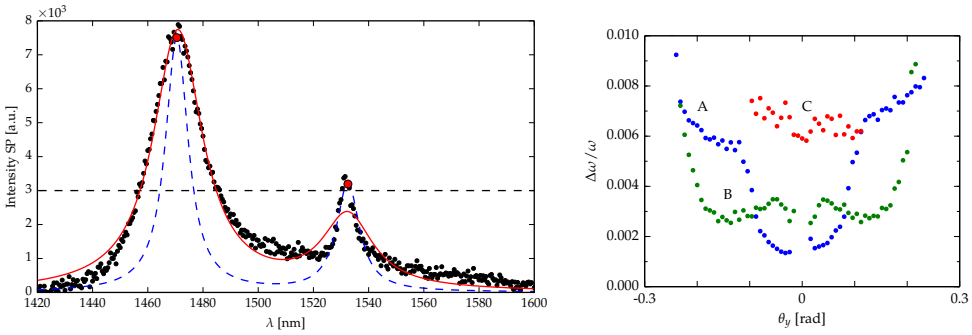
Figure A.1: *Fit example on the dataset used in chapter 4 at $\theta_y = -0.12$ radian. Our code first counts the number of peaks, indicated by the red dots, which are above the threshold of $0.2{\cdot}10^4$ as indicated by the black dashed line. The blue dashed line then shows the initial guess, based on the initial parameters chosen as explained in subsection A.1.3. The code varies the parameters of this initial curve to match the data and then returns the fit parameters. The red solid line shows the final fit.*

A.2 Quality of the fits

Overall the fits on the data were very good. We wrote an interactive program in iPython notebook to scan through the fits, in this way we could easily see which fits were off. Figure A.2 is an example of a fit gone wrong. When fits went wrong, we would simply delete the fitted parameters out of the parameter plots.

We experienced different places were fits would go wrong. If a peak was found at relatively low SNR, it was very hard to find the correct fit. This peak could easily be omitted by raising the threshold. When peaks became too sharp and the data was dominated by another peak (which is the case in figure A.2), the fits also tended to go wrong. And finally, at the edge of our range we also experienced bad fits. This is mainly because we are at the edge of the bandgap or Fermi level, where the Lorentzian peaks sometimes are a bit misshaped. Also, the SNR is often very low at these points.

All of these problems almost always resulted in higher widths than ex-



(a) **(b)**

Figure A.2: ***(a)*** *Fit example of a 470 nm array (not the one of our main results) at $\theta_y = -0.23$ rad, which went wrong. We observe a sharp peak from the B band which is dominated by a big peak from the A band. This is the reason why the fit went wrong. This and other similar fits are deleted by hand. The fitted width of the peak is much larger than our data shows.* ***(b)*** *Spectral widths parameter plot of the same array. Near $\theta_y = -0.23$ rad the fitted spectral widths of the B band were found to increse suddenly, which triggered us to take a look at the fits.*

pected. For this reason we used the plot of the fitted spectral widths as an indication of how good our fits were. If the spectral widths would increase suddenly, we would check the quality of the fits. Most of the time the fits were indeed bad, but not always. The spectral widths of the A band in figure 4.4 also suddenly increased near the edge, but the fits were actually still very accurate.

Appendix B

Loss and scattering of surface plasmons on optically-pumped hole arrays

We performed a detailed study on the $a_0 = 470$ nm array from Frerik's sample, as used in chapter 4. Also, the model used in ref. [5] has been expanded to the coupled-mode model as mentioned in chapter 3. These two subjects are highlighted in our article currently under review at the Journal of Optics, entitled "Loss and scattering of surface plasmons on optically-pumped hole arrays". The version as submitted in june 2014 is included in this appendix.

Loss and scattering of surface plasmons on optically-pumped hole arrays

V T Tenner, A N van Delft, M J A de Dood, M P van Exter

Huygens-Kamerlingh Onnes laboratorium, Leiden University, The Netherlands

E-mail: tenner@physics.leidenuniv.nl

Abstract.

We study surface plasmons on 2-dimensional square arrays of sub-wavelength holes in a gold film deposited on an optically-excited semiconductor. We observe four resonances of which we measure the central frequencies, the spectral widths, and the relative intensities. The spectral widths allow us to quantify various loss processes, including ohmic loss, optical absorption/gain and radiative scattering loss. We distinguish between two bright/radiative modes and two dark/non-radiative modes. Prominent kinks in the plasmon dispersion relation are attributed to Rayleigh anomalies, associated with light lines in the semiconductor. We are able to model most of these observations, but challenges remain.

Keywords: surface plasmon, dispersion relation, metal hole array

Submitted to: *J. Opt.*

1. Introduction

Surface plasmons (SPs) are combined optical and electronic solutions to Maxwell equations. Such surface plasmons can be localized on nano particles, or they can be traveling along an extended interface. One remarkable property of SPs is the strong confinement normal to the interface, with exponentially decaying fields in both media. The associated decay length is typically sub-wavelength and depends on the dielectric constant of the materials. Several applications benefit from the strong confinement of the field. This confinement can be used to enhance the light-matter interaction, for example in molecular sensors [1, 2]. The strong confinement can also be used to guide SP and build plasmonic circuits, including beam splitters [3] and phase manipulators [4]. Furthermore, meta-materials often consist of sub-wavelength metal structures. The plasmon modes living on these structures are responsible for many of their extraordinary properties, such as the ability to create negative refractive index materials [5, 6] and primitive cloaking devices [7].

The limiting factor in many SP-based system is the SP loss, which can be divided in (ohmic) absorption loss and scattering loss. Even SPs traveling on smooth Ag or Au interfaces are typically absorbed after 2-100 μm [8] for wavelengths between 0.5 and 1.5 μm. Gain can be introduced to overcome this limitation and several demonstrations of full loss compensation have been shown, using a semiconductor [9] or dye [10] as gain material. SP-lasing action is possible when full loss compensation is combined with optical feedback. Different feedback mechanisms have been demonstrated in several structures [11], including metal-coated nano-pillars [12], gold nano-spheres [13], and metal hole arrays [14].

In this article, we study metal hole arrays with gain below their lasing threshold. The dispersion relation of plasmonic bands in a metal hole arrays with gain have been observed before [15], but until now, no detailed study of the spectral widths of these bands existed. We present the first systematic study of the central frequency, the spectral width, and the intensity of these bands. The observed widths allow us to quantify the effective ohmic loss as well as the scattering loss rate. We describe the observed complex band structure with a coupled-mode model that includes only four SP modes.

2. Methods

2.1. Sample

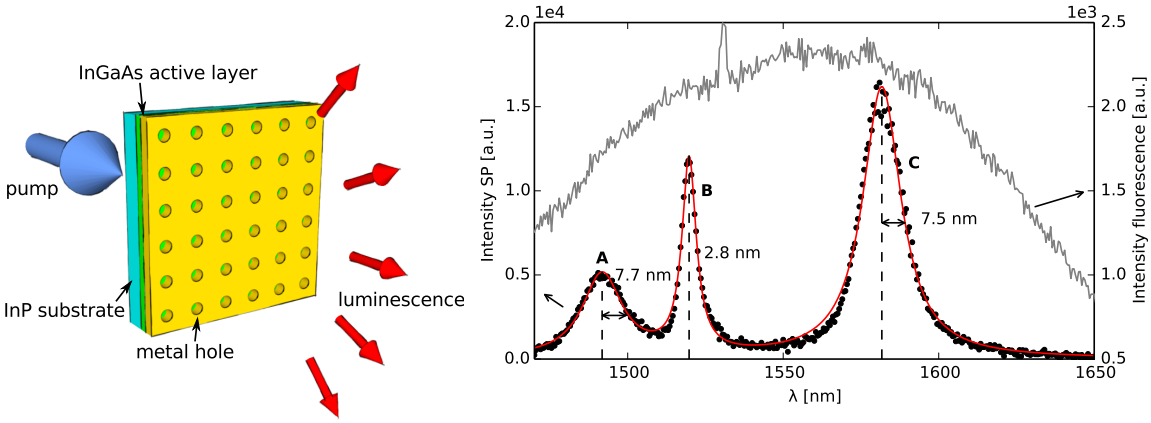


Figure 1. a) Sketch of setup, with an optical pump through the substrate and fluorescence through the hole array. b) Spectrum of p-polarized light emitted at ($\theta_x = 0$, $\theta_y = 0.10$) rad. The fit (red curve) is based on a sum of Lorentzian resonances. The half widths of three peaks are indicated in the graph. The grey line shows the direct fluorescence light, measured in absence of the gold layer.

The semiconductor-gold samples that we studied are identical to the ones described in refs [14, 15]. Their layer structures, depicted in Fig. 5, comprise a 100 nm thick gold film on a 105 nm thick $In_xGa_{1-x}As$ semiconductor layer that is lattice matched ($x \simeq 0.53$) to a 300 μm thick, double polished, InP substrate. A 15-20 nm thick SiN_x/InP spacer-layer between the $In_xGa_{1-x}As$ and the gold prevents quenching of the optically excited semiconductor. A 20 nm thick chromium layer on top of the gold damps the SP resonance on the gold-air interface, leaving only the gold-semiconductor resonances. The gold layer is perforated with holes with a diameter of 180 nm. The holes are arranged in a 50x50 μm square grid with a spacing $a_0 = 470$ nm between the holes.

2.2. Experimental geometry

Fig. 1a shows the experimental geometry. We optically excited the InGaAs active/gain layer through the InP substrate with a cw pump laser with a wavelength of 1064 nm. The incident power (120 mW) was below the laser threshold of the device at a temperature of 120 K. The pump spot on the sample had a square shape with a size of 45 μm, created by passing a Gaussian laser beam through a hologram. This pump beam excites electron-hole pairs in the active layer. When these hot carriers recombine radiatively, their energy is converted into photons or surface plasmons living on the interface between the gold and the InGaAs semiconductor layer. The grey curve in Fig. 1b shows the measured spectrum of this radiative recombination for a part of the sample without gold.

The collection part of our setup was the same as in [14, 15]. The fluorescence light was collected in the far field at the air/vacuum side of the sample as a function of angle (θ_x, θ_y). More precisely, the light was collected using a 20x microscope objective with a numerical aperture of 0.4.

Sequentially the array structure was magnified 20x by a lens with a 20 cm focal length. Finally, the far field of the sample was reached in the back-focal plane of a lens (focal length 5 cm), where the fluorescence light was collected on a spectrometer using a single mode optical fiber ($w \approx 10\mu$m @$\lambda = 1500$ nm). This fiber was scanned in the far field by a 2-dimensional translation stage. In this article we keep the fiber at $\theta_x = 0$ and scan the fiber only in the θ_y direction, from $\theta_y = -0.44$ rad to $\theta_y = 0.44$ rad in 110 steps, but we limit all figures to the most relevant part $-0.3 \leq \theta_y \leq 0.3$. The angle θ_y is related to the in-plane momenta in the y-direction (Γ-M) by $k_{\|} = \omega/c \sin\theta_y$ [15]. A thin film polarizer was used to discriminate between vertical (p=TM) and horizontal (s=TE) polarized light.

Figure 1b shows a typical p-polarized emission spectrum, obtained at an angle $\theta_y = 0.10$. This spectrum contains three peaks, labeled by A, B and C. A curve consisting of (the incoherent sum of) three Lorentzians fits the data very well. We repeated these measurements at other angles θ_y and for both polarizations, where similar high quality fits yielded the position, the spectral width (half height half maximum, HWHM), and the area of each of the three peaks. The observed angle dependence of the frequencies, spectral widths, and areas of these resonances are discussed below.

3. Measured data

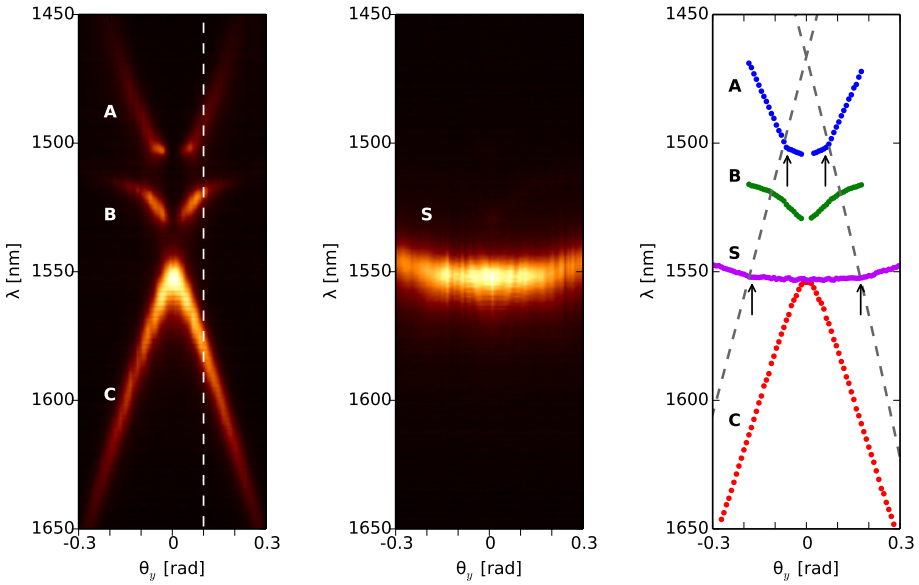


Figure 2. False color plots of the measured fluorescence intensity $I(\theta_y, \lambda)$ of a) p-polarization and b) s-polarization. The y-axis depicts energy, in units of vacuum wavelength of the collected photons. The white dashed line in a) at $\theta_y = 0.10$ indicates the cross section shown in Fig. 1. c) The central frequencies of the bands deduced from the measurements. The dashed lines show the light line of the semiconductor, corresponding to the Rayleigh anomaly. The arrows mark clear kinks in the dispersion relation.

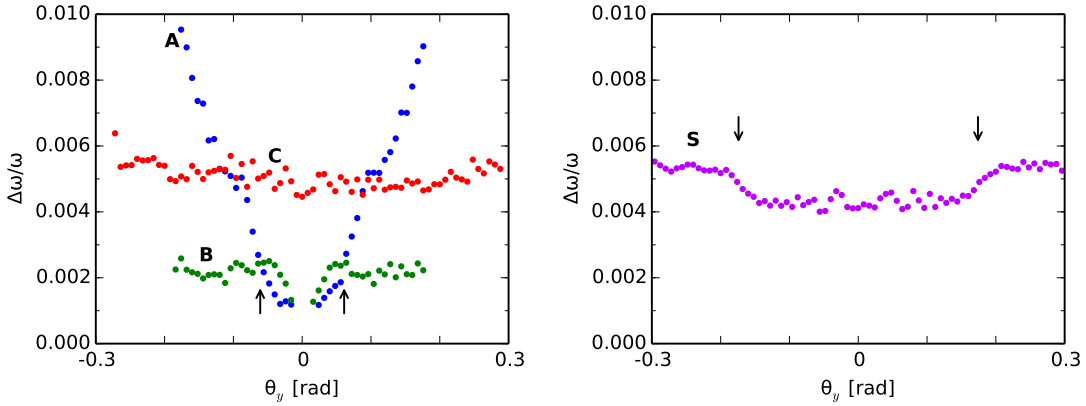


Figure 3. Fitted linewidths of the a) p-polarized b) s-polarized bands. The angles for the observed kinks in the A and S band are marked by the arrows.

Measurements of the p- and s-polarization are shown in Fig. 2a and b. In Fig. 2a three p-polarized resonances are visible, labeled with A, B and C. Fig. 2b shows one intense s-polarized resonance, labeled S.

3.1. Measurement of resonant frequencies

Figure 2c shows the resonance frequencies of all bands. The central frequency of the s-polarized S band is remarkably constant up till $\theta = \pm 0.17$ and then increases mildly towards larger θ. The p-polarized A, B, and C bands are more dispersive. At large angle $|\theta| > 0.1$ the A and C band show the approximately linear dispersion expected for traveling-wave SPs, but at smaller angles these bands exhibit an avoided crossing, again with a remarkably flat segment in the A band. This crossing is modified by the presence of the (intermediate) B band, which inherits a much stronger dispersion for small angles [15].

Sharp kinks in the A and S band at $\theta = \pm 0.06$ and $\theta = \pm 0.17$, respectively, are marked with arrows in Fig. 2c. The A and S band are more flat between the kinks. The B and C band do not show kinks.

3.2. Measurement of linewidth

Figure 3 shows the observed linewidths of the four SP bands. The C and S band have a linewidth of $\Delta\omega/\omega \simeq 0.0045$ near $\theta = 0$, while the A and B band have a smaller linewidth of $\Delta\omega/\omega \simeq 0.001$. The C band is the only band with a linewidth that does not depend on angle. The linewidth of the S band is almost constant, except from the sudden increase at $\theta = \pm 0.17$, marked with arrows in Fig. 3b. The linewidth of the B band is M shaped, with a minimum at $\theta = 0$. The linewidth of the A band also has a minimum at $\theta = 0$, but its linewidth increases rapidly towards larger angles.

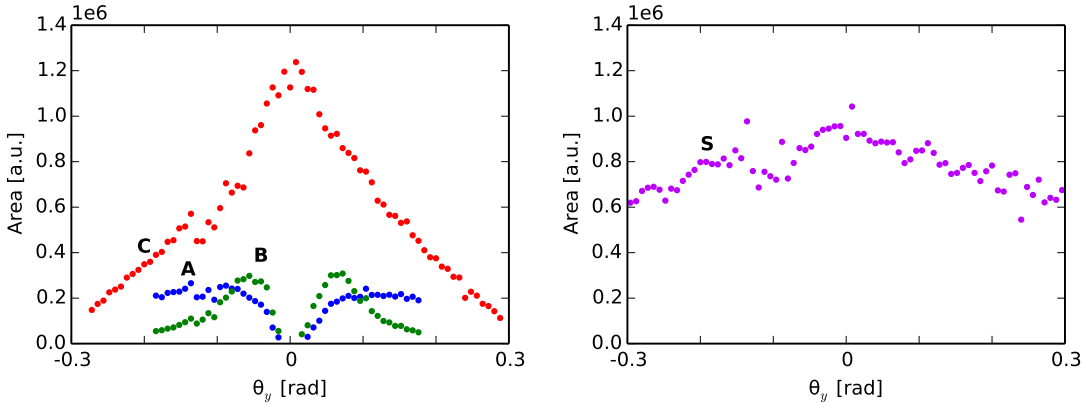


Figure 4. Integrated intensity of a) p-polarized and b) s-polarized bands. Note the large area of the bright C and S mode and the small area of the dark A and B mode, in particular at $\theta = 0$.

3.3. Measurement of intensity/area

Figure 4 shows the integrated intensity $A = \pi I_{max} \Delta\lambda$ of each resonance emitted on the air/vacuum side, where I_{max} is the maximal intensity of the band and $\Delta\lambda \ll \lambda$ is its spectral width (HWHM). The C and S band have a large area, whereas the A and B band are much fainter. The C and S band are brightest at $\theta_y = 0$, whereas the A and B band have almost no light at $\theta_y = 0$ and acquire intensity only at $\theta_y \neq 0$. We call the C and S mode bright/radiative and the A and B mode dark/non-radiative [15].

3.4. Measurement on different sample

We repeated these measurements and analysis on different metal hole arrays on the same wafer, with different hole sizes and spacings ($a_0 = 450...470$ nm). All arrays show similar results. To be more precise: (i) the resonance frequency of the S band had kinks again, (ii) the spectral width of the two bright modes C and S was larger than that of the two dark modes A and B. The spectral width of the S band suddenly increased between $\theta_y = 0.1$ and 0.25, and the spectral width of the B band was M-shaped, (iii) the C and S bands were bright, while the A and B bands were dark at $\theta_y = 0$. There were also differences, one of them was in the width of the A band, which was constant at small angles, but increased towards larger angles. The onset of this increase is at the same angle as the kink in the resonance frequency of the A band.

4. Discussion

We analyzed all our data with a model based on four coupled surface plasmon modes. This model, which is discussed in the Appendix, extends the work of [15]. It includes the SP-to-SP scattering that dominantly determines the SP dispersion relation [15]. As an extension it also includes SP losses due to SP-to-photon scattering, ohmic losses, and optical gain. These loss/gain channels determine the spectral widths of the resonances. We describe these losses as imaginary rates in the

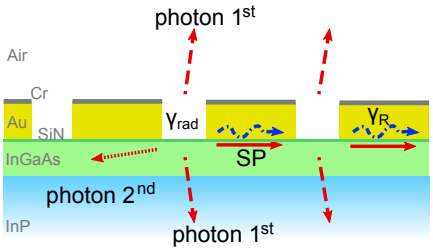


Figure 5. Loss channels of SP. The reduced ohmic loss, denoted by the loss rate γ_R, is spatial uniform and includes the effective optical gain. The radiative loss, denotes by the loss rate γ_{rad}, occurs only at the holes and originates from scattering of SPs to photons. Under some conditions, the SP-to-photon scattering is distributed over three diffraction orders: 1st order diffraction into the air, and 1st and 2nd order diffraction into the semiconductor. We only observe the emission into air and analyze it as a function of frequency and emission angle.

coupled mode model and thus obtain a complex band structure.

Our model contains the following parameters: SP-to-SP coupling is parameterized by scattering rates γ and κ, which refer to back-scattering and right-angle-scattering, respectively [15]. SP-to-photon coupling occurs at a radiative loss rate γ_{rad}. Finally, the ohmic loss and optical gain are combined into a reduced ohmic loss rate $\gamma_R = \gamma_r - g(\omega)$, where the ohmic loss rate γ_r includes surface-roughness related effects. Note that the optical gain $g(\omega)$ is frequency dependent and determined by the parabolic electronic bands of the semiconductor and their Fermi-Dirac type spectral filling. This bandfilling is also visible in the spontaneous emission spectrum of the bare sample depicted in Fig. 1b.

The various contributions to the loss rate are depicted in Fig. 5. The radiative loss originates from scattering of surface plasmons to photons in air and to photons in the semiconductor. In the semiconductor a second scattering channel opens up at sufficiently large angle θ. This so called Rayleigh anomaly is expected to lead to an increase in the radiative loss [16]. This second scattering channel is denoted as photon 2nd in Fig. 5, where the other channels are denoted as photon 1st.

4.1. Analysis of resonance frequencies

The position of the resonance frequencies can be described reasonably well by the coupled mode model, as explained in the Appendix. This model predicts three p-polarized bands - in the form of two hyperbolic bands A and C and one V shaped band B - and a single weakly hyperbolic s-polarized band S [15].

Our measured dispersion curves can be well fitted with amplitude scattering rates $\gamma/\omega_0 = +0.012$ and $\kappa/\omega_0 = \pm 0.004$ (see Appendix). These values are comparable to values found in ref. [15], where SP dispersion curves measured for nine similar arrays could be well fitted with rates $\gamma/\omega_0 = +0.013$ to $+0.017$ and $\kappa/\omega_0 = \pm 0.005$ to ± 0.011.

Our relatively simple model does not predict the kinks and flat region of the A and S band. We attribute these effects to the so-called Rayleigh anomaly depicted in Fig. 5, which results in additional losses at large angles and might even affect the SP dispersion via a Kramer-Krönig-type relation. The dashed light lines in Fig. 2c correspond to the Rayleigh anomaly for a substrate index $n \simeq 3.12$, which is close to the value of $n = 3.10$ of the InP wafer.

4.2. Analysis of linewidths

The coupled mode theory described in the Appendix explains most of the features that we observe. For instance, it explains why the SPs in the C and S band are more lossy than the SPs in the A and B band. The former are radiative and scatter efficiently to photons, whereas the latter are non-radiative and do not scatter to photons at small angles [17]. At $\theta = 0$ the spectral width of the radiative modes $\gamma_{C,S} = 2\gamma_{rad} + \gamma_R$ is set by the radiative and reduced ohmic losses, whereas the spectral width of the dark modes $\gamma_{A,B} = \gamma_R$ is only determined by the reduced ohmic losses. The experimental results presented in Fig. 3 correspond to $\gamma_R/\omega \leq 0.001$ and $2\gamma_{rad}/\omega \simeq 0.004$. The coupled mode model also predicts the M shape of the B band with the smallest width at $\theta = 0$ (see Appendix). The Rayleigh anomaly explains why the SP loss in the S-band increases suddenly at the arrows.

Which loss rates and refractive index do we expect theoretically? For surface plasmons on a smooth metal-semiconductor interface we have

$$n_{eff} + in_i = \sqrt{\frac{\epsilon_1\epsilon_2}{\epsilon_1 + \epsilon_2}} \tag{1}$$

Insertion of $\epsilon_2 = -116 + 11.1i$ for gold at $\lambda \approx 1.5$ μm and $\epsilon_1 = 11.6$ ($n_1 = 3.407$) for transparent InGaAs, yields $n_{eff} + in_i \approx 3.588 + 0.019i$. Inclusion of the absorption loss/gain via $\epsilon_1 = 11.6 \pm 0.65i$, yields $n_{eff} + in_i \approx 3.587 + 0.130i$ and $n_{eff} + in_i \approx 3.591 - 0.093i$, respectively. As the experiment is performed on a thin layer of InGaAs only, we need to include the influence of the InP substrate, such that the actual SP index will lie between the quoted numbers and the value $n_{eff} + in_i \approx 3.235 + 0.013i$ calculated for SPs on a gold-InP interface ($\epsilon_1 \approx 9.61$ as $n_1 \approx 3.10$ for InP).

The expression $\gamma_r/\omega = n_i/n_{eff}$ yields an expected ohmic loss rate of $\gamma_r/\omega \approx 0.005$ for transparent InGaAs. Because the observed reduced ohmic loss rate $\gamma_R/\omega = 0.001$ is considerably lower than the loss rate of transparent InGaAs, the InGaAs layer has a clear gain component.

In ref. [9], Van Beijnum et al. determined the loss rate of the radiative SP mode around $\theta = 0$, using Fano fits of measured transmission spectra. Their analysis yielded a half-width (HWHM) around $\lambda = 1492$ nm of 24 nm at low pump power and 4 nm at high pump power. The corresponding loss rates are $\gamma_s/\omega = 0.016$ and 0.0027, respectively. The loss rate of the unpumped system is considerably larger than the loss rates observed here, while the loss rate of the pumped system is somewhat smaller. For comparison, we expect a reduced ohmic loss rate of $\gamma_R/\omega \approx 0.036$ for the most absorptive case of unpumped bulk InGaAs.

4.3. Analysis of areas

The first thing to note in Fig. 3b is the distinction between the bright C and S and the dark bands A and B. This observation is easily understood from the theoretical observation that the bright bands have field maxima at the holes, whereas the dark bands have field minima at the holes around $\theta = 0$. The field we are referring to is the parallel magnetic field, which is the field that scatters SPs to photons [15].

We also note that the area of the B band has the expected M-shape. The area of A and C band, however, do not behave as expected but rather decreases at high angles θ. We attribute this decrease to the spectral shape of the fluorescent light of the optically-pumped gain layer, which is the light-source in our experiment. This spectral shape is shown in Fig. 1b. This spontaneous emission is only emitted at frequencies above the bandgap of the semiconductor, but below the

filling of the electronic band of the semiconductor. At high angles θ the resonance frequencies of the A and C band are located at the edge of the fluorescence spectrum and hence less light couples into these bands.

5. Conclusion

We studied the complex dispersion relation of surface plasmons on a square metal hole array. We observed the frequencies, linewidths and areas of the four resonances and found remarkable features, which can be attributed to a Rayleigh anomaly. Furthermore, we made a distinction between bright (radiative) and dark (non-radiative) modes. From the linewidth we deduced scattering rates for the photon and ohmic losses. The SP-to-photon scattering losses are approximately three times higher then the other losses SP in the system.

In order to gain more understanding of the measured complex SP dispersion relation, we developed a coupled mode model that only includes four modes. This model predicts several observed features, but some features are not yet fully understood. Numerical simulations could provide some insight in the remaining questions. More important, we need to increase the complexity of the model by adding the three physical phenomena, that are mentioned at the end of the Appendix: (i) the frequency dependence of spontaneous emission and gain, (ii) the frequency dependence of all scattering rates, and (iii) the occurrence of a Rayleigh anomaly at large angles.

Acknowledgments

We acknowledge Frerik van Beijnum, Peter J. van Veldhoven and Erik Jan Geluk for their help in fabricating the samples at the COBRA Research Institute of the Technische Universiteit Eindhoven, The Netherlands.

Appendix A. Appendix for 'Loss, gain and scattering ...' article

Appendix A.1. SP field in traveling-wave basis

In this appendix we describe a simple coupled-mode model for the surface plasmons (SP) that exist at the metal-dielectric interface of a square hole array. This model is an extension of an earlier model described in ref. [15]. It is based on the notion that photons emitted with parallel momentum $\mathbf{k}_{\parallel} \equiv (k_x, k_y)$ only couple to traveling-wave SPs with (in-plane) momenta $\mathbf{k}_{SP} = \mathbf{k}_{\parallel} + \mathbf{G_i}$, where $\mathbf{G}_i$ is a lattice vector of the hole array. Our model only considers the four dominant Fourier components (modes), with magnitude $|\mathbf{G}_i| \equiv G = (2\pi/a_0)$ pointing in either of the four lattice directions $\{\mathbf{e}_x, \mathbf{e}_{-x}, \mathbf{e}_y, \mathbf{e}_{-y}\}$, and neglects higher-order Fourier components of the SP. These components are not resonant and hardly excited. The non-resonant contributions from all Fourier components are included in a change in the effective refractive index n_{eff}, which is different from the effective index of SPs on a smooth interface.

In this article, we only consider emission at $\theta_x = 0$ and write the parallel momentum $k_y = (\omega/c)\sin\theta_y$ at emission angles $\theta_y \ll 1$. We express the associated out-of-plane components of the SP field at position $\mathbf{r} \equiv (x, y)$ as

$$\mathbf{E}(\mathbf{r}, t) = \left[E_x(t)\mathbf{u}_x e^{iGx} + E_{-x}(t)\mathbf{u}_{-x}e^{-iGx} + E_y(t)\mathbf{u}_y e^{iGy} + E_{-y}(t)\mathbf{u}_{-y}e^{-iGy}\right] e^{ik_{\parallel}y}, \quad \text{(A.1)}$$

where $\{E_x, E_{-x}, E_y, E_{-y}\}$ are the modal amplitudes of the four traveling waves. The Eigenvectors $\mathbf{u}_i$, which describe the four associated optical polarizations, are chosen to be rotationally-imaged

copies of each other. When the four modal amplitudes in Eq. (A.1) are combined into a single vector $|E\rangle$, the time evolution of this SP field can be expressed as $d|E\rangle/dt = -iH|E\rangle$, where H is a 4×4 matrix. If scattering is neglected, H reduces to a diagonal matrix with elements $\{\omega_{+x}, \omega_{-x}, \omega_{+y}, \omega_{-y}\}$. At $\theta_x = 0$ and $\theta_y \equiv \theta$, the dispersion relation $\omega = |\mathbf{k}_\parallel|c/n_{eff}$ for traveling-wave SPs on a metal-dielectric interface yields $\omega_{\pm y}(\theta) \approx (G \pm k_\parallel)c/n_{eff} \approx \omega_0 \pm c_1\theta$, and $\omega_x(\theta) = \omega_{-x}(\theta) \approx \omega_0 + c_2\theta^2$, with $\omega_0 \equiv (2\pi/a_0)c/n_{eff}$, $c_1 \equiv \omega_0/n_{eff}$, and $c_2 \equiv \omega_0/(2n_{eff}^2)$, where n_{eff} is the SP effective index.

Reference [15] also described the SP-to-SP scattering processes. This scattering was divided into three fundamental processes: forward scattering under 0° (at a rate γ_0), right-angle scattering under ±90° (at a rate κ), and backwards scattering under 180° (at a rate γ). Inclusion of these scattering processes into the $d|E\rangle/dt = -iH|E\rangle$ matrix description yields the result presented in ref. [15]:

$$H = \begin{pmatrix} \omega_0 + c_2\theta^2 & \gamma & \kappa & \kappa \\ \gamma & \omega_0 + c_2\theta^2 & \kappa & \kappa \\ \kappa & \kappa & \omega_0 + c_1\theta & \gamma \\ \kappa & \kappa & \gamma & \omega_0 - c_1\theta \end{pmatrix} \tag{A.2}$$

The scattering rates γ and κ are assumed to be real-valued, such that the associated coupling is conservative (= energy conserving) and the matrix H is Hermitian. The reference frequency is again $\omega_0 = (2\pi/a_0)c/n_{eff}$, but n_{eff} now contains a small contribution from forward scattering at a rate γ_0.

Appendix A.2. SP field in standing-wave basis

Next, we transform the evolution matrix H from the traveling-wave to the standing-wave basis. For this purpose, we combine the waves traveling in the $\pm x$ direction into two standing waves with out-of-plane $E_\perp(\mathbf{r}) = E_{cx}(\mathbf{r}) \propto \cos Gx \cdot \exp ik_y y$ and $E_{sx}(\mathbf{r}) \propto \sin Gx \cdot \exp ik_y y$, where the labels cx and sx denote a cosine- or sine-pattern in the x direction. Likewise, we combine the $\pm y$ traveling waves into two standing waves with out-of-plane E-fields $E_{cy}(\mathbf{r}) \propto \cos Gy \cdot \exp ik_y y$ and $E_{sy}(\mathbf{r}) \propto \sin Gy \cdot \exp ik_y y$. The transition from the $\{+x, -x, +y, -y\}$ traveling-wave basis to the $\{i \sin\{Gx\}, \cos Gx, i \sin Gy, \cos Gy, \}$ or $\{sx, cx, sy, cy\}$ standing-wave basis transforms the H matrix into

$$H = \begin{pmatrix} \omega_0 + c_2\theta^2 - \gamma & 0 & 0 & 0 \\ 0 & \omega_0 + c_2\theta^2 + \gamma & 0 & 2\kappa \\ 0 & 0 & \omega_0 - \gamma & \delta \\ 0 & 2\kappa & \delta & \omega_0 + \gamma \end{pmatrix}, \tag{A.3}$$

where $\delta \equiv -c_1\theta = -(\omega_0/n_{eff})\theta$.

Note how the 4×4 matrix separates in an uncoupled element, associated with the sx standing wave, and a 3×3 matrix. This separation results from the mirror symmetry in the xz-plane ($\theta_x = 0$). The $E_{sx}(\mathbf{r}) \propto \sin Gx \cdot \exp ik_y y$ field is the only standing wave that is odd under mirror inversion, while the other three standing waves are even. The former couples only to s-polarized emission, which has an odd symmetry, while the latter three mix and couple to p-polarized emission, which also has an even symmetry. Also note how the coupling rate κ, associated with SP-SP scattering under 90°, only couples the $\cos Gx$ and $\cos Gy$ waves, which have intensity maxima at the holes, while the detuning δ only couples the cy and sy standing waves.

Appendix A.3. Losses and gain

Next, we include losses and gain in our model. First of all, we include ohmic losses at an ohmic damping rate $\gamma_r = (n_i/n_{eff})\omega$, where n_i and n_{eff} are the imaginary and real part of the SP effective index $n_{eff} + in_i = \sqrt{\epsilon_1\epsilon_2/(\epsilon_1+\epsilon_2)}$ at the interface between medium 1 and 2. Second, we include optical gain at a gain rate g_0 per second. As both ohmic loss and optical gain are distributed approximately uniformly, they can be combined into an effective ohmic loss rate $\gamma_R \equiv \gamma_r - g_0$ and are easily incorporated in our matrix description by replacing the evolution matrix $-iH \rightarrow -iH - \gamma_R$. Finally, we include radiative losses through scattering from SPs to photons, either in the air or in the semiconductor. This radiative loss rate $\gamma_{rad} = \gamma_{air} + \gamma_{semi}$ is equal for all traveling waves, but mode selective for the standing waves. As radiative scattering only occurs at the holes and as the SP-to-photon coupling proceeds dominantly via the in-plane magnetic field (at $\theta \ll 1$), we expect that only the standing waves with a sinusoidal $E_\perp$- pattern couple radiatively, at a decay rate $2\gamma_{rad}$, while the cosine-type modes don't couple. By combining the above loss and gain rate into our matrix description we arrive at our final expression

$$H = \begin{pmatrix} \tilde{\omega}_{sx} & 0 & 0 & 0 \\ 0 & \tilde{\omega}_{cx} & 0 & 2\kappa \\ 0 & 0 & \tilde{\omega}_{sy} & \delta \\ 0 & 2\kappa & \delta & \tilde{\omega}_{cy} \end{pmatrix} , \tag{A.4}$$

where $\tilde{\omega}_i = \omega_i + i\gamma_i$ are four complex frequencies, with real parts $\omega_{sx} = \omega_0 + c_2\theta^2 - \gamma$, $\omega_{cx} = \omega_0 + c_2\theta^2 + \gamma$, $\omega_{sy} = \omega_0 - \gamma$, and $\omega_{cy} = \omega_0 + \gamma$, and imaginary parts $\gamma_{sx} = \gamma_{sy} = \gamma_R + 2\gamma_{rad}$ and $\gamma_{cx} = \gamma_{cy} = \gamma_R$.

Appendix A.4. Spontaneous emission spectra

The fluorescence spectrum of the optically-pumped systems can be calculated in two steps, once the evolution matrix of the SP-field is known. In the first step, the spontaneous emission into the surface plasmon manifold is calculated from the expression

$$\frac{d}{dt}|E(t)\rangle = -iH|E(t)\rangle + |S\,(t)\rangle \quad \Rightarrow \quad |E\,(\omega)\rangle = -i(H-\omega)^{-1}|S\,(\omega)\rangle\,,$$

where the 4-element vector $|S\rangle$ describes the original spontaneous emission, divided over the four standing-wave SP modes, and $|E\rangle$ describes the generated SP field. The multiplication by $(H-\omega)^{-1}$ describes how the original emission source is modified by the gain and loss in the system into the resulting SP field, which therefore peaks around optical frequencies ω close to the complex poles of the matrix H. In the first step of the calculation, we assume that the original emission at each optical frequency is equally distributed over the four standing waves and that the four emitted fields are uncorrelated. This assumption is the Fourier equivalent of the statement that the original emission at different spatial positions is homogeneous, isotropic, and uncorrelated. In the second (and final) step, the generated surface plasmons are scattered into photons and detected. The *sx* standing SP wave scatters into *s*-polarized photons and only the *sy* standing SP wave is assumed to scatter into *p*-polarized photons.

The emission in the (odd) *s*-polarized mode is easily calculated. As only one of the four SP standing waves is odd, the SP field follows from the scalar relation $E\,(\omega) = iS\,(\omega)\,/(\omega - \tilde{\omega}_{sx})$ and its absolute square

$$I_{\mathrm{s-polarization}}\,(\omega) \propto \frac{I_s\,(\omega)}{(\omega-\omega_{sx})^2 + \gamma_{sx}^2}\,, \tag{A.5}$$

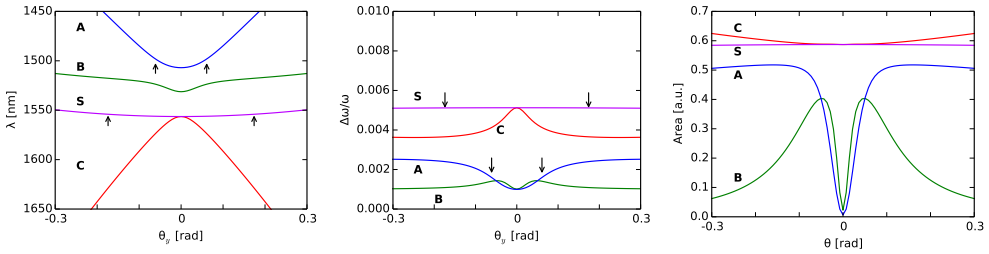


Figure A1. Numerical results for four SP bands (A, B, C, and S) obtained from a simple coupled-mode model: a) central frequency (= dispersion relation), b) linewidth, and c) spectral area. The arrows indicate the angle at which a Rayleigh anomaly is observed in the experiment.

where $I_s(\omega) \leftrightarrow |S(\omega)|^2$. We recognize the standard Lorentzian form, with its resonance frequency $\omega_{sx} \equiv \omega_0 + c\delta^2 - \gamma$ and its (HWHM) half-width $\gamma_{sx} \equiv \gamma_R + 2\gamma_{rad}$. The excited sx mode emits efficiently from SP-to-photon, as this $E_{sx}(\mathbf{r}) \propto \sin Gx \cdot \exp ik_y y$ mode has a magnetic field $\mathbf{H}_{sx}(\mathbf{r}) \propto \cos Gx \cdot \exp ik_y y \mathbf{e}_y$ with anti-nodes at the holes.

A calculation of the (even) p-polarized emission is more complicated, as this emission originates from three coupled SP modes. After some straightforward mathematics, which involves the inversion of a 3×3 matrix and a projection onto the sy mode, which is the only one of the three mode that couples to photons, we obtain

$$I_{\mathrm{p-polarization}}(\omega) \propto \sum_{i=1}^{3} \frac{I_s(\omega)\,|\langle sy|u_i\rangle|^2}{(\omega-\omega_i)^2+\gamma_i^2}\ , \tag{A.6}$$

where $\langle u_i|$ are the three left eigenvectors of the 3×3 lower-right submatrix $\tilde{H}$ of H, $\tilde{\omega}_i \equiv \omega_i + i\gamma_i$ are the associated eigenvalues, such that $\langle u_i|\tilde{H} = \tilde{\omega}_i\langle u_i|$, and $|sy\rangle$ denotes the field of the sy mode. The explicit solution of this problem reads

$$I_{\mathrm{p-polarization}}(\omega) \propto \frac{|\,(\tilde{\omega}_{cy}-\omega)(\tilde{\omega}_{cx}-\omega)-4\kappa^2|^2+\delta^2\left[|\tilde{\omega}_{cx}-\omega|^2+4\kappa^2\right]}{|\,(\tilde{\omega}_{sy}-\omega)\left[(\tilde{\omega}_{cy}-\omega)(\tilde{\omega}_{cx}-\omega)-4\kappa^2\right]-\delta^2(\tilde{\omega}_{cx}-\omega)\,|^2}\ I_s(\omega)\ , \tag{A.7}$$

At $\delta = 0$ and $\kappa = 0$, we recover the expected result $I_{\mathrm{p-polarization}}(\omega) = I_{\mathrm{s-polarization}}(\omega) \propto I_s(\omega)/|\tilde{\omega}_{sy}-\omega|^2$.

The above expressions for the emitted spectrum contain subtleties that might go unnoticed in their present form, as they are related to the frequency dependence of some of its parameters. The spontaneous emission spectrum $I_s(\omega)$ and the associated stimulated emission or optical gain $g(\omega)$ will for instance depend on frequency. Although their spectral dependence is relatively smooth and broad in semiconductors, it is something that should be included in a more complete model. Even the SP-to-SP scattering rates γ and κ and the SP-to-photon scattering rate γ_{rad} are expected show a strong wavelength dependence, with Rayleigh-type scattering rates scaling as ω^n with $n \geq 6$, depending on the type of scattering [18]. A third and final point to include in a more complete model is the sudden increase in radiative loss at the appearance of a new diffraction order (Rayleigh anomaly).

Appendix A.5. Comparison to measurement

In this section we present a prediction of our model. Figure A1 shows the theoretical dispersion relation, linewidth and the area of the four SP bands for the experimentally observed angles. These theortical curves are calculated with the experimentally observed parameters, namely: $\gamma/\omega = 0.012$, $\kappa/\omega = 0.004$, $\gamma_{rad}/\omega = 0.002$, $\gamma_R = 0.001$.

Figure A1(a) shows a theoretical calculation of the central frequencies of the four SP bands. The model predicts one s-polarized band, the S band, with a weakly hyperbolic shape. Furthermore, the model predicts three p-polarized bands: two hyperbolic bands, the A and C band, and a V shaped B band. Note that the S and C band are degenerate at $\theta = 0$. The angles at which there are kinks in the experimental dispersion relation are indicated by the arrows in Fig. A1a.

Figure A1(b) shows the theoretical linewidths of the four bands. The model predicts two broad modes, the bright C and S band, and two narrow modes, the dark A and B band. The width of the C and S band are both equal to $2\gamma_{rad} + \gamma_R$ at $\theta = 0$. The width of the S band is constant, whereas the width of the C band decreases somewhat towards larger angles. This decrease is not observed in our measurement. The linewidths of the dark A and B band is small, in particular at $\theta = 0$ where they are both equal to γ_R. At larger angles, the linewidth of the A band gradually increases, whereas the angle dependence of the linewidth of the B band is M shaped. These generally shapes are observed experimentally, but the linewidth of the A band at large angles exceeds the width of the C band in our experiment.

Figure A1(c) shows the theoretical spectral area of the four bands. Around $\theta = 0$, we again distinguish the bright C and S band form the dark A and B band, which acquire intensity only at $\theta \neq 0$. Apart from these generic features, there is no clear match between the theoretically calculated and experimentally measured spectral areas. The main reason for this mismatch is the frequency dependence of the spontaneous emission, which in our simple model was assumed to be frequency independent (see main text).

References

[1] Alexandre G. Brolo, Reuven Gordon, Brian Leathem, and Karen L. Kavanagh. Surface plasmon sensor based on the enhanced light transmission through arrays of nanoholes in gold films. *Langmuir*, 20(12):4813–4815, May 2004. 2

[2] Ahmet A. Yanik, Min Huang, Osami Kamohara, Alp Artar, Thomas W. Geisbert, John H. Connor, and Hatice Altug. An optofluidic nanoplasmonic biosensor for direct detection of live viruses from biological media. *Nano Lett.*, 10(12):4962–4969, November 2010. 2

[3] H. Ditlbacher, J. R. Krenn, G. Schider, A. Leitner, and F. R. Aussenegg. Two-dimensional optics with surface plasmon polaritons. *Applied Physics Letters*, 81(10):1762–1764, September 2002. 2

[4] A. Melikyan, L. Alloatti, A. Muslija, D. Hillerkuss, P. C. Schindler, J. Li, R. Palmer, D. Korn, S. Muehlbrandt, D. Van Thourhout, B. Chen, R. Dinu, M. Sommer, C. Koos, M. Kohl, W. Freude, and J. Leuthold. High-speed plasmonic phase modulators. *Nat Photon*, 8(3):229–233, March 2014. 2

[5] R. A. Shelby, D. R. Smith, and S. Schultz. Experimental verification of a negative index of refraction. *Science*, 292(5514):77–79, April 2001. 2

[6] D. R. Smith, J. B. Pendry, and M. C. K. Wiltshire. Metamaterials and negative refractive index. *Science*, 305(5685):788–792, August 2004. 2

[7] Tolga Ergin, Nicolas Stenger, Patrice Brenner, John B. Pendry, and Martin Wegener. Three-dimensional invisibility cloak at optical wavelengths. *Science*, 328(5976):337–339, April 2010. 2

[8] William L. Barnes, Alain Dereux, and Thomas W. Ebbesen. Surface plasmon subwavelength optics. *Nature*, 424(6950):824–830, 2003. 2

[9] Frerik van Beijnum, Peter J. van Veldhoven, Erik Jan Geluk, Gert W. 't Hooft, and Martin P. van Exter. Loss compensation of extraordinary optical transmission. *Applied Physics Letters*, 104(6):061112, February 2014. 2, 8

[10] Israel De Leon and Pierre Berini. Amplification of long-range surface plasmons by a dipolar gain medium. *Nat Photon*, 4(6):382–387, June 2010. 2

[11] Ren-Min Ma, Rupert F. Oulton, Volker J. Sorger, and Xiang Zhang. Plasmon lasers: coherent light source at molecular scales: Plasmon lasers: coherent light source at molecular scales. *Laser & Photonics Reviews*, 7(1):1–21, January 2013. 2

[12] Martin T. Hill, Yok-Siang Oei, Barry Smalbrugge, Youcai Zhu, Tjibbe de Vries, Peter J. van Veldhoven, Frank W. M. van Otten, Tom J. Eijkemans, Jaroslaw P. Turkiewicz, Huug de Waardt, Erik J. Geluk, Soon-Hong Kwon, Yong-Hee Lee, Richard Notzel, and Meint K. Smit. Lasing in metallic-coated nanocavities. *Nature Photonics*, 1(10):589–594, September 2007. 2

[13] M. A. Noginov, G. Zhu, A. M. Belgrave, R. Bakker, V. M. Shalaev, E. E. Narimanov, S. Stout, E. Herz, T. Suteewong, and U. Wiesner. Demonstration of a spaser-based nanolaser. *Nature*, 460(7259):1110–1112, August 2009. 2

[14] Frerik van Beijnum, Peter J. van Veldhoven, Erik Jan Geluk, Michiel J. A. de Dood, Gert W. 't Hooft, and Martin P. van Exter. Surface plasmon lasing observed in metal hole arrays. *Phys. Rev. Lett.*, 110(20):206802, May 2013. 2, 3

[15] M. P. van Exter, V. T. Tenner, F. van Beijnum, M. J. A. de Dood, P. J. van Veldhoven, E. J. Geluk, and G. W. 't Hooft. Surface plasmon dispersion in metal hole array lasers. *Optics Express*, 21(22):27422, November 2013. 2, 3, 4, 5, 6, 7, 8, 9, 10

[16] H. Gao, J. M. McMahon, M. H. Lee, J. Henzie, S. K. Gray, G. C. Schatz, and T. W. Odom. Rayleigh anomaly-surface plasmon polariton resonances in palladium and gold subwavelength hole arrays. *Optics Express*, 17(4):2334, February 2009. 7

[17] Z. L. Cao and H. C. Ong. Determination of the absorption and radiative decay rates of dark and bright plasmonic modes. *Optics Express*, 22(13):16112, June 2014. 8

[18] F. van Beijnum, A.S. Meeussen, C. Rétif, and M.P. van Exter. Rayleigh scattering of surface plasmons by sub-wavelength holes. *Optics Express*, 22(9):10317, May 2014. 12