EXPERIMENTAL AND NUMERICAL STUDY OF
PLASMONIC-ENHANCED BULK AND SURFACE
NONLINEAR OPTICAL EFFECTS

Nuno Vitor da Silva Braz

A dissertation submitted for the degree of Doctor of Philosophy
Department of Electronic and Electrical Engineering
University College London

November 18, 2016

I, Nuno Vitor da Silva Braz confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Abstract

Metallic nanostructures of dimensions below the wavelength of light show remarkable optical properties due to their ability to trap light through excitation of free-electron oscillations at a metal-dielectric interface. Coined localized surface plasmons, these are associated with a strong electro-magnetic field localization in the vicinity of the nanostructures and can find applications in integrated nonlinear optical systems given the favourable power-law scaling of nonlinear effects with the electric-field. The enhancement of surface and bulk nonlinear optical effects were investigated in this work using a variety of polarization-sensitive plasmonic devices.

Investigation of bulk second- and third-order nonlinear effects in isotropic non-centrosymmetric materials were realised, both experimentally and numerically, using electric-field enhancements from complementary asymmetric cross metasurfaces. Devices were fabricated using ZnO as a nonlinear material, containing linear plasmonic resonances in the near- to mid-infrared range fully characterized using Fourier transform infra-red spectroscopy. Nonlinear optical measurements have been conducted on the complementary samples, and a tunable nonlinear optical response was observed through control of the plasmonic modes by changing the incident polarization or wavelength. A multipole expansion of the linear and nonlinear polarizations show that electric dipole interactions dominate the optical response of complementary samples. However, as result of variations in the spatial distribution of the electric-fields, predicted by Babinet’s principle, significant differences in the magnetic-dipole and electric-quadrupole response between complementary samples have been observed. In particular, the device made of cross apertures in a metallic film shows the strongest enhancement of the two samples. The miniaturization, con-
trol and optimization of nonlinear optical effects are of interest for future integrated photonic systems.

A numerical study of surface second-order nonlinear effects from chiral metamaterials has been also performed, showing strong correlation between regions of superchiral light and enhanced nonlinear effects. Furthermore, a device structure has been realised with a non-reciprocal nonlinear response dependent on the handedness of the incident light. Numerical calculations based on a multipole expansion of the second-harmonic polarization, show that variations in electric quadrupolar contributions are responsible for the dichroic response.
Acknowledgements

The present work would have not been possible without the guidance and support of several people. I would like to express my sincere appreciation to my supervisor Prof. Paul A. Warburton for providing the resources and guidance to much of the experimental work carried out in this thesis. He has enabled me to pursue my own research interests, and has taught me the value of a critical mind in the analysis of scientific work. I would also like to express my gratitude to my secondary supervisor Prof. Nicolae C. Panoiu, which has led me the way into to the fast growing field of nonlinear plasmonics, and forged a number of opportunities that shaped the course of this research.

I am hugely thankful to Felicitas Walter and Prof. Thomas Zentgraf for welcoming me at University of Paderborn (Germany). They have enabled me the use of their optical facilities, and provided the much needed support and knowledge in performing the nonlinear optical measurements presented in this thesis. I am also grateful to Dr. Guy Matmon and Prof. Paul F. McMillan from UCL for the use of their laboratory equipment and facilities, which were instrumental for the linear characterization of my devices. I would like to thank Steve Etienne, Suguo Huo, Rohit Khanna, Vj Krishnan, Lorella Rossi and Richard Thorogate from LCN for their training and friendly support in all matters relating to cleanroom fabrication and microscopy tools. I also thank Dr. Ventsislav Valev from University of Bath for allowing me to be part of a fruitful collaboration which initiated me on the modelling of nonlinear plasmonic effects. Dr. Edward Osley for initiating me in the project, including training and maintenance advice on the FTIR. I am also very grateful to Maddison Coke for the MBE growth and XRD characterization of ZnO films used
in this work. A very special thank you also to Dr. Arnaud Blois for proof reading my thesis and for providing very constructive comments.

I gratefully acknowledge financial student support by the Engineering and Physical Sciences Research Council, as part of the centre for doctoral training for photonic systems development (grant EP/G037256/1).
List of Acronyms

2DEG  two-dimensional electron gas

Al$_2$O$_3$  sapphire

CCD  charge coupled device

CD  circular dichroism

CPL  circularly polarized light

EBL  electron-beam lithography

EM  electromagnetic

EOT  extraordinary optical transmission

FIB  focused ion-beam

FTIR  Fourier transform infra-red spectroscopy

GaAs  gallium arsenide

HSQ  hydrogen silsesquioxane

IPA  isopropyl alcohol

IR  infrared

LSP  localized surface plasmon

MBE  molecular beam epitaxy

MCT  mercury cadmium telluride
MIBK  methyl isobutyl ketone
OPO  optical parametric oscillator
PEC  perfect electric conductor
PMMA  poly(methyl methacrylate)
PMT  photon multiplier tube
RCWA  rigorous coupled wave analysis
SAM  spin angular momentum
SEM  scanning electron microscope
SHG  second-harmonic generation
SP  surface plasmon
SPP  surface plasmon polariton
THG  third-harmonic generation
TE  transverse electric
TM  transverse magnetic
UV  ultra violet
XRD  X-ray diffraction
ZnO  zinc oxide
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^{(n)}$</td>
<td>n-th order nonlinear susceptibility</td>
</tr>
<tr>
<td>$\Delta k$</td>
<td>wave vector mismatch</td>
</tr>
<tr>
<td>$\hbar$</td>
<td>reduced Planck's constant $[\frac{\hbar}{2\pi} = 1.054571800 \times 10^{-34} \text{ J s}]$</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>extinction coefficient</td>
</tr>
<tr>
<td>$L$</td>
<td>effective path length [m]</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>free space wavelength [m]</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>vacuum permeability $[4\pi \times 10^{-7} \text{ H/m}]$</td>
</tr>
<tr>
<td>$\omega$</td>
<td>angular frequency [rad/s]</td>
</tr>
<tr>
<td>$\tilde{\nu}$</td>
<td>spectroscopic wavenumber [m$^{-1}$]</td>
</tr>
<tr>
<td>$\tilde{n}$</td>
<td>complex refractive index, $\tilde{n} = n + i\kappa$</td>
</tr>
<tr>
<td>$\varepsilon_0$</td>
<td>vacuum permittivity $[(c^2\mu_0)^{-1} = 8.854187817... \times 10^{-12} \text{ F/m}]$</td>
</tr>
<tr>
<td>$\varepsilon_r$</td>
<td>relative permittivity</td>
</tr>
<tr>
<td>$c$</td>
<td>speed of light in vacuum $[299792458 \text{ m/s}]$</td>
</tr>
<tr>
<td>$e$</td>
<td>elementary charge $[1.602176565 \times 10^{-19} \text{ C}]$</td>
</tr>
<tr>
<td>$h$</td>
<td>Planck’s constant $[6.626070040 \times 10^{-34} \text{ J s}]$</td>
</tr>
<tr>
<td>$k/k$</td>
<td>wavenumber/wavevector [m]</td>
</tr>
<tr>
<td>$m^*$</td>
<td>effective mass</td>
</tr>
</tbody>
</table>
\( m_e \quad \text{electron mass } [9.109 \, 382 \, 91 \times 10^{-31} \, \text{kg}] \)

\( n \quad \text{refractive index} \)

\( n_e \quad \text{free electron density } [\text{m}^{-3}] \)

\( n_{e,o} \quad \text{extraordinary/ordinary refractive index} \)

\( Z_0 \quad \text{vacuum impedance } [c\mu_0 = 119.916 \, 983 \, 2\pi \, \Omega] \)
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.7</td>
<td>Fabrication processes for plasmonic metamaterials</td>
<td>57</td>
</tr>
<tr>
<td>3.2</td>
<td>Optical characterization</td>
<td>60</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Fourier transform infra-red spectroscopy</td>
<td>60</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Light amplification by stimulated emission of radiation (laser)</td>
<td>66</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Optical parametric oscillator</td>
<td>69</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Monochromator</td>
<td>71</td>
</tr>
<tr>
<td>3.2.5</td>
<td>Optical detection systems</td>
<td>72</td>
</tr>
<tr>
<td>3.2.6</td>
<td>Measurement of nonlinear frequency generation</td>
<td>74</td>
</tr>
<tr>
<td>3.3</td>
<td>Modelling methods</td>
<td>81</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Rigorous coupled wave analysis</td>
<td>82</td>
</tr>
<tr>
<td>4</td>
<td>Tunable bulk nonlinear effects using complementary metamaterials</td>
<td>87</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>87</td>
</tr>
<tr>
<td>4.2</td>
<td>Device description</td>
<td>89</td>
</tr>
<tr>
<td>4.3</td>
<td>Fabrication and ZnO film characterization</td>
<td>91</td>
</tr>
<tr>
<td>4.4</td>
<td>Numerical methods</td>
<td>95</td>
</tr>
<tr>
<td>4.5</td>
<td>Linear optical characterization</td>
<td>97</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Transmission and reflection measurements on complementary samples</td>
<td>97</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Babinet’s principle and field enhancement</td>
<td>102</td>
</tr>
<tr>
<td>4.5.3</td>
<td>Effect of the ZnO film on the LSP resonances</td>
<td>105</td>
</tr>
<tr>
<td>4.5.4</td>
<td>Effect of scaling the unit cell on the LSP resonances</td>
<td>106</td>
</tr>
<tr>
<td>4.5.5</td>
<td>Output polarization state</td>
<td>107</td>
</tr>
<tr>
<td>4.6</td>
<td>Nonlinear optical characterization</td>
<td>110</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Experimental methods</td>
<td>110</td>
</tr>
<tr>
<td>4.6.2</td>
<td>Power dependence</td>
<td>111</td>
</tr>
<tr>
<td>4.6.3</td>
<td>Reference sample measurements</td>
<td>114</td>
</tr>
<tr>
<td>4.6.4</td>
<td>Polarization dependence</td>
<td>116</td>
</tr>
<tr>
<td>4.6.5</td>
<td>Wavelength dependence</td>
<td>123</td>
</tr>
<tr>
<td>4.7</td>
<td>Conclusion</td>
<td>132</td>
</tr>
<tr>
<td>5</td>
<td>SHG-CD response from superchiral metasurfaces</td>
<td>134</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>134</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.2</td>
<td>Gammadion L-shaped devices with tunable SHG-CD response</td>
<td>136</td>
</tr>
<tr>
<td>5.3</td>
<td>Experimental methods and results</td>
<td>138</td>
</tr>
<tr>
<td>5.4</td>
<td>Numerical calculations</td>
<td>141</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Linear optical response</td>
<td>141</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Orbital angular momentum flux</td>
<td>146</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Optical chirality</td>
<td>149</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Nonlinear optical response</td>
<td>156</td>
</tr>
<tr>
<td>5.5</td>
<td>Conclusions</td>
<td>159</td>
</tr>
<tr>
<td>6</td>
<td>Conclusions</td>
<td>161</td>
</tr>
<tr>
<td>6.1</td>
<td>Future work</td>
<td>165</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Plasma angular frequency $\omega_p$ for different metals, showing resonance in the ultra violet range. Frequency calculated using Eq. 2.5 with free electron densities from Ashcroft and Mermin [29], and plasma wavelength from $\lambda = \frac{2\pi c}{\omega_p}$ where $c$ is the speed of light. .......................... 9

2.2 Dispersive permittivity of gold. Individual contributions from the free-electrons (intraband, Drude model) and bound-electron (interband, Lorentz model) transitions are shown, including sum/total contributions [31]. Experimental data was taken from Palik [32]. .......... 10

2.3 (a) Band structure and (b) total density of states for bulk gold calculated using GPAW software [33]. The arrows highlight regions of interband transitions. The peaks in the density of states come from the 5d bands. (c) Reflectivity of gold obtained from the expression $R = \left| \frac{\tilde{n} - 1}{\tilde{n} + 1} \right|^2$ where $\tilde{n}$ is the complex refractive index converted from the relative permittivity of gold obtained from Rakic [31]. .............. 12

2.4 Dispersion curve for the free electron gas and free propagating light in vacuum. ............................................................ 14

2.5 Field distribution of an surface plasmon (SP) propagating as a p-polarized or TM wave, i.e. no magnetic field allowed in the direction of propagation. ............................................................... 16

2.6 Dispersion curve of an SP, bulk plasmon and free propagating light using two different dielectric materials, vacuum ($\varepsilon_r^d = 1$) and CaF$_2$ ($\varepsilon_r^d = 2$). The Drude model for ideal metals (no losses, $\gamma = 0$) was considered in the description of the dispersive relative permittivity $\varepsilon_r^m(\omega)$ (Eq. 2.14). ................................. 19
2.7 Calculation of the SP propagation length $L_x$ and z-decay length inside a dielectric $L_x^d$ and a metal $L_x^m$. The mediums on each half of the interface are vacuum ($\varepsilon_r^d = 1$) and gold. The Lorentz-Drude model in Eq. 2.7 was used to calculate the relative permittivity $\varepsilon_r^m$ of gold.

2.8 Dispersion curves for SPP using the (a) prism and (b) grating coupling methods. For the optical prism case, phase matching occurs for incident angles $\arcsin(\sqrt{\varepsilon_g/\varepsilon_p}) \leq \theta_i \leq \pi/2$ where evanescent waves and SPP wave vectors are equal.

2.9 Plot of the absolute and phase values of the polarizability of a gold sphere embedded in vacuum (solid line, $\varepsilon_d = 1$) and water (dashed line, $\varepsilon_d = 1.8$), calculated using Eq. 2.22. Dispersive permittivity of gold was obtained from the Lorentz-Drude model in Eq. 2.7 using the parameters in Rakic et al. [31]. Inset shows a sketch of the gold sphere subject to an electric-field.

2.10 Scattering cross section calculated using Eq.2.23 for a 50 nm gold sphere in a water based solution ($\varepsilon_d = 1.8$). The color of the dispersed gold sphere solution is determined by the difference in absorbed (not shown) and scattered cross sections.

2.11 Incident and transmitted EM-fields through complementary objects.

2.12 Nonlinear optical coefficients for IR materials with (a) second- and (b) third-order nonlinearities. Values obtained from [108, 109, 110, 107, 111].

2.13 Coherence length for (a) GaAs and (b) ZnO materials using different incident pump wavelengths. The refractive index values used in Eq. 2.29 were obtained from [112] and [113]. The different curves in (b) correspond to the ordinary and extraordinary refractive indexes of ZnO.

2.14 Refractive index of GaAs showing regions of normal and anomalous dispersion [32]. Note the peaks in the extinction coefficient $\kappa$ - related with the absorption coefficient by $\alpha = 4\pi\kappa/\lambda$ - are associated with the regions of anomalous dispersion.
List of Figures

3.1 Schematic cross-section of a Carl Zeiss Gemini electron column (adapted from Carl Zeiss’s application manual).............................................. 44

3.2 SEM micrographs of various resist structures based on the EBL parameters presented in Tab. 3.1................................................................. 47

3.3 FTIR transmission measurement of an array of plasmonic slit apertures fabricated using the EBL at 10 kV acceleration voltage, 30 µm aperture, and varied electron dosage (73 to 165 μC/cm²). Structures were milled using argon milling for 15 min at V_{acc}=390 V, I_{b}=10 mA and V_{b}=200 V. Substrate material is CaF₂, and thickness of metallic films 30 nm gold and 5 nm titanium. FTIR measurement using linear polarized light, with electric-field vector orthogonal to the long dimension of the slits. The images in the dose labels correspond to a SEM micrograph of a single slit, fabricated with the specified dosage. The dashed line shows the simulated result obtained using a commercial implementation of RCWA (Sec. 3.3.1). The model was implemented using N=8 harmonics and a fixed refractive index value of 1.39 for the CaF₂ substrate................................................. 48

3.4 (a) FTIR transmission measurement of an array of plasmonic slit apertures, (b) transmission versus dose plot for three wavelengths, corresponding to features in the transmission data. Samples were fabricated using a gallium FIB at 30 kV acceleration voltage, 50 pA beam current, and varied ion dose (0.01 to 0.1 nC/cm²). Substrate material is CaF₂, and thickness of metallic films 30 nm gold and 5 nm titanium. FTIR measurement using linearly polarized light, with electric-field vector orthogonal to the long dimension of the slits................................................................. 50

3.5 Schematic diagram of a simplified version of an argon ion-miller. Image adapted from [146]................................................................. 51
### List of Figures

3.6 FTIR transmission measurement with linear polarized light, of an array of plasmonic slit apertures fabricated using the EBL at 10 kV with a 30 μm aperture and a dose of 128 μC/cm², and varied milling times at $V_{acc}$=390 V, $I_b$=10 mA and $V_b$=200 V. Substrate material is CaF₂, and thickness of the films 30 nm gold and 5 nm titanium. Incident electric-field polarization with vector orthogonal to the long dimension of the slits. ............................... 53

3.7 Schematic diagram of a resistive and electron beam evaporator. Adapted from lecture notes by Prof. Andrew Flewitt at University of Cambridge. 55

3.8 Sample process flow for two plasmonic configurations based on a range of fabrication techniques. In processes (B,G) the plasmonic devices are etched into the substrate, while for processes (A,C-F) the nanostructures are sitting on the surface. ....................... 58

3.9 Schematic representation of an FTIR system including both transmission and reflection measurement modes. ....................... 62

3.10 (a) Single-sided, bi-directional interferogram and (b) Fourier transformed single beam spectrum of an Al₂O₃ substrate. Interferogram obtained using a resolution of 4 cm⁻¹ and 800 scans. Interferogram was phase corrected and spectrum calculated using the Mertz method. Forward and backward scans were transformed and phase corrected separately, and combined at the end using a mean function. .... 62

3.11 Transmission of a ZnO film on a Al₂O₃ substrate, normalized with respect to a background measurement of the substrate. ............... 64

3.12 Energy levels in a four level laser system. .............................. 67

3.13 Absorption and emission spectra of a Ti:sapphire (Al₂O₃) rod. Data extracted from [151]. ............................... 69

3.14 Power spectra from a Chameleon™ Compact OPO-Vis OPOs pumped by the Chameleon™ Ultra II laser. Data extracted from official manufacturer system specification. ............... 70
3.15 (a) Schematic diagram of an unfolded Czerny-Turner monochromator and (b) spectral efficiency of a 150 lines/mm and 500 nm blaze wavelength grating. Adapted from Andor Shamrock 303i specification manual.

3.16 Spectral detectivity (D*) of various visible and IR detectors. This is a measure of the signal-to-noise performance of a detector, and generally a large value means good detectivity. The MCT data was obtained from the manufacturers test report (InfraRed Associates, Inc. FTIR-7-1.0), the pyroelectric detector data from Hamamatsu [153], and the PMT and CCD data from Newport [154].

3.17 Optical setup for nonlinear transmission measurements.

3.18 Relative laser intensity for different $\theta_1$ and $\theta_2$ polarizer angles according to Malus’s Law. Angle $\theta_2$ can be used to control the polarization incident on the sample with some compromise on the intensity of the light. Red line highlights the optimum combination of $\theta_1$ and $\theta_2$ which gives the highest laser intensity.

3.19 Schematic diagram of a confocal optical microscope used in the surface SHG measurements in Chap. 5. In the current configuration the microscope is operating in reflection mode.

3.20 (a,b) SEM micrograph of two plasmonic samples exposed to 125TW/m$^2$ and 320TW/m$^2$ intensities from a Nd:YAG pico second laser, and (c) FTIR response of sample (b) before and after nonlinear measurements. The samples consist of an array of slit apertures on a gold film filled with GaAs. The dimensions of the slits are as described in the inset of picture (c). The plasmonic response is characterized by a single localized surface plasmon (LSP) resonance at 8µm, when excited with an electric-field perpendicular to the longer arm of the slit.

3.21 Example geometry of a 2D grating (adapted from Moharam [159]).
3.22 Illustration of the Gibbs phenomenon on the near-field distribution at two different interfaces: (a,b) a dielectric medium and (c,d) at the surface of a metal with complex permittivity. The model consists of a gold cross-shaped metallic structure on a Al$_2$O$_3$ substrate and covered with a zinc oxide (ZnO) dielectric medium. The simulation was performed using an electric-field polarization parallel to the vertical arm of the cross, and wavelength matching the LSP resonance of the vertical arm. The near field distributions were calculated using N=16 harmonics and represent a z-cut across a (left) dielectric interface and a (right) metallic interface. Images (b,d) correspond to a x-cut across the plane highlighted by the dashed lines in figures (a,c) for three different simulations using N=[8, 16, 24] harmonics.

4.1 Helium focused ion beam micrograph of an array of (a) transmissive and (b) reflective cross samples. The mean dimensions of the crosses in both samples are $W_V=220$ nm, $W_H=70$ nm, $L_H=620$ nm, $L_V=380$ nm and $P=760$ nm, with labels specified in Fig. 4.2. The illustrations and material labels serve only for differentiating between plasmonic devices. The micrograph samples are not gold or sapphire terminated but have a ZnO film on top of the metallic nanostructures.

4.2 (a) top and (b) side view of an asymmetric cross geometry, showing main cross dimensions, in-plane electric field polarization angle $\psi$, k-vector angle of incidence $\theta$, and azimuthal angle $\varphi$.

4.3 Process steps for the fabrication of an array of asymmetric (a) cross apertures and (b) opaque crosses using EBL, argon-ion milling and MBE.

4.4 (a) X-ray diffraction (XRD) measurement of transmissive and reflective samples showing peak at $2\theta/\omega=34.4^\circ$ corresponding to polycrystalline ZnO film with (0002) orientation. (b) Pole figure XRD measurement on both samples showing preferred orientation of the polycrystalline grains with respect to the nanostructures.
4.5 (a,b) transmission and (c,d) reflection measurements for the (a,c) transmissive and (b,d) reflective samples using a FTIR. Measurement is achieved at normal incidence $\theta = 0^\circ$ with the electric field polarisation angle varying from $\psi=0^\circ$ (black) to $\psi=90^\circ$ (purple) with a $10^\circ$ step. Resolution $4 \text{ cm}^{-1}$ and 800 scans used in both background and sample measurements. Dashed vertical lines highlight position of LSP resonances.

4.6 Numerically calculated (total) transmitted power from the (a) transmissive and (b) reflective samples at fundamental wavelength, normal incidence $\theta = 0^\circ$ and varying incident polarization $\psi=0^\circ$ to $90^\circ$ with a $10^\circ$ step. Dashed vertical lines highlight position of LSP resonances. Plots (c,d,e) and (f,g,h) show the electric dipole, magnetic dipole and electric quadrupole contributions, for the transmissive and reflective samples respectively. Individual multipole contributions normalized with respect to the peak power $S(\omega)$.

4.7 Numerically calculated transmitted linear power from individual and total multipole contributions for the (a) transmissive and (b) reflective samples at resonant wavelengths $\lambda^T_1=1866 \text{ nm}$ and $\lambda^R_1=1828 \text{ nm}$ respectively. Source was set at normal incidence $\theta=0^\circ$ and data normalized with respect to maximum total power intensity.

4.8 Numerically calculated (total) linear transmitted power from the (a) transmissive and (b) reflective samples for wavelengths corresponding to the SHG and THG of the fundamental range in Fig. 4.6. Simulation was performed at normal incidence $\theta = 0^\circ$ and using incident polarizations $\psi = [0^\circ, 90^\circ]$. Results were normalized with respect to the maximum total power in Fig. 4.6.
4.9 Simulated normalized electromagnetic (EM)-fields across the (a-d) transmissive and (e-g) reflective crosses when excited at plasmonic resonances $\lambda_1$ and $\lambda_2$ with $\psi = 0^\circ$ and $\psi = 90^\circ$ polarizations (respectively). Top and bottom rows show electric- and magnetic-field profiles (respectively) for the xy-plane at the gold-ZnO interface ($z= 55$ nm in Fig. 4.10). The black arrows give the corresponding $|E|$ and $|B|$ field directions. ................................................................. 103

4.10 Cross section of the linear polarization $|P(\omega, r)|$ across the (a-b) transmissive and (c-d) reflective crosses when excited at plasmonic resonances $\lambda_1$ and $\lambda_2$ with $\psi = 0^\circ$ and $\psi = 90^\circ$ polarizations respectively. The black arrows give the corresponding $|P(\omega)|$ linear polarization direction. ................................................................. 103

4.11 Reflection measurement at normal incidence $\theta=0^\circ$ on reflective sample before ZnO growth. Resolution $4\text{cm}^{-1}$ and 800 scans used in the measurement. ................................................................. 106

4.12 LSP resonances $\lambda^R_1$ and $\lambda^R_2$ for the reflective sample, without the ZnO film, as function of the scaling of the cross-shaped unit cell. Measurements at normal incidence $\theta=0^\circ$. .................................................. 107
4.13 Transmission (top) and reflection (bottom) measurements using a polarizer and analyzer configuration. Each column represent a measurement of one of seven incident polarizations $\psi$, for a range of wavelengths and 29 analyzer angles. The analyzer angle is defined with respect to the x-axis of the nanostructures. Top and bottom rows correspond to data for a transmissive and reflective sample (respectively). FTIR measurements are slightly noisier when operating in reflection, and for that reason, we have used a reflective sample with a larger unit cell to improve SNR. A factor of $1.875 \times$ was used on all unit cell dimensions of the reflective sample, to red-shift the LSP resonances towards a measurable range. Dashed white lines highlight resonant peaks, i.e. $\lambda^T_1=1888$ nm and $\lambda^T_2=2667$ nm for the transmissive sample, and $\lambda^R_2=2960$ nm and $\lambda^R_2=3770$ nm for the reflective sample. Resolution of $16\text{cm}^{-1}$ and 200 scans were used in the background and sample measurement.

4.14 Schematics of the experimental setup for measuring bulk nonlinear harmonic generation from the complementary plasmonic samples. A half-retarder and polarizer are used to control the power and incident electric field polarisation from a femto-second laser, spectrally tunable using an optical parametric oscillator system. The collimated beam is then focused onto the nanostructures, from the unstructured side of the sample, and a long pass filter is used to remove unwanted nonlinearities from the optical system. The transmitted light, containing both fundamental and harmonic generated signals, is then captured by an objective and coupled into a spectograph, which spatially separates photons of different energies and records these by using an air-cooled Si-CCD detector.
4.15 SHG and THG measurements as function of the incident power and at normal incidence $\theta=0^\circ$. (a) corresponds to a measurement of the transmissive sample using incident polarization $\psi = 0^\circ$ and excitation wavelength $\lambda_T^T = 1888 \text{nm}$, and (b) corresponds to the reflective sample with $\psi = 90^\circ$ and $\lambda_R^R = 1816 \text{nm}$. (c,f) Power dependence measurements of the (green) SHG and (blue) THG signals from the transmissive and reflective sample respectively. Data points were obtained from the peak intensity values in plots (a,b,d,e) and power meter data. 

4.16 SHG and third-harmonic generation (THG) measurement on two reference samples of (a) transmissive and (b) reflective nanostructures using three different incident polarizations $\psi=[0^\circ, 45^\circ, 90^\circ]$. Both samples were excited at normal incidence $\theta=0^\circ$ with a 100 mW average power and fixed wavelength- $\lambda_T^T=1879 \text{nm}$ for the transmissive sample and $\lambda_R^R=1884 \text{nm}$ for the reflective sample. Data collected over 9 accumulations of 5 s exposures.

4.17 Nonlinear contributions from (black) bare Al$_2$O$_3$ substrate and (red) ZnO film deposited on Al$_2$O$_3$ - no nanostructures on either of the samples. Measurement was achieved at normal incidence $\theta=0^\circ$ with a 75 mW average power, 1868 nm incident wavelength and fixed arbitrary polarization $\psi$. Data collected over 12 accumulations of 5 s exposures.
4.18 (a,d) SHG and (b,e) THG measurements as function of incident electric-field polarization $\psi$ for the (a,b,c) transmissive and (d,e,f) reflective samples. At normal incidence $\theta = 0^\circ$ and with linearly polarized light, devices were excited at plasmonic resonance $\lambda_1$ and with a 100mW average power. Plots (c) and (f) were extrapolated from the peak values obtained from plots (a,b) and (d,e), and Fourier transform infrared spectroscopy (FTIR) data in Fig. 4.5.(a) and 4.5.(d). Solid lines correspond to the power $S(\omega_g)$ obtained from Eq. 4.3 using a multipole expansion of the linear, SHG and THG polarizations (Eq. 4.1a, 4.1b and 4.1c) for excitation of the plasmonic mode $\lambda_1$ (as defined in Fig. 4.6) and varying incident electric-field polarization $\psi$. Dashed red line corresponds to a fit of the function $A\sin^2 \psi + B\cos^2 \psi$ with $A=1$ and $B=0.19$ for transmissive sample, and $A=1$ and $B=0.13$ for the reflective sample.

4.19 Total lattice power and separate multipole contributions as function of incident polarization $\psi$, from the (left) transmissive and (right) reflective samples for (top) SHG and (bottom) THG nonlinear processes. In the model, the source was set to normal incidence $\theta=0^\circ$, and wavelengths $\lambda_T^1=1866$ nm and $\lambda_R^1=1828$ nm for the transmissive and reflective samples (respectively). Data normalized to the maximum total power of each individual nonlinear process and device.

4.20 Spectral measurement of the (a,c) linear (solid lines) and (b,d) SHG intensity (dots) for three different incident polarizations $\psi = [0^\circ, 45^\circ, 90^\circ]$, normal incidence $\theta=0^\circ$ and pulse power 60mW. Left column corresponds to data from the transmissive sample, and right the reflective sample. Dashed lines are the results from the numerical calculations.

4.21 Complete numerical spectral analysis of (a,d) linear and (b,c,e,f) SHG effects from the (left) transmissive and (right) reflective samples for a range of incident polarizations $\psi$. Bottom row corresponds to the SHG power after accounting for the linear optical response of the metastructures at SHG wavelengths, as shown in Fig. ??.
4.22 Multipole decomposition of the SHG process from the (left) transmissive and (right) reflective samples, and for a range of incident polarizations $\psi$. Top, middle and bottom plots correspond to SHG power contributions from the electric-dipole $S_{ed}(2\omega)$, magnetic-dipole $S_{md}(2\omega)$ and electric-quadrupole $S_{eq}(2\omega)$ respectively, normalized to the total radiated power $S(2\omega)$ in Fig. 4.21.(b) and Fig. 4.21.(e). 128

4.23 Spectral measurement of the (a,c) linear (solid lines) and (b,d) THG intensity (dots) for three different incident polarizations $\psi = [0^\circ, 45^\circ, 90^\circ]$, normal incidence $\theta=0^\circ$ and pulse power 60mW. Left column corresponds to data from the transmissive sample, and right the reflective sample. Dashed lines are the results from the numerical calculations. 130

4.24 Complete numerical spectral analysis of (a,d) linear and (b,c,e,f) THG effects from the (left) transmissive and (right) reflective samples for a range of incident polarizations $\psi$. Bottom row corresponds to the THG power after accounting for the linear optical response of the metastructures at THG wavelengths, as shown in Fig. 4.8. 131

5.1 Schematic depiction of the L-shaped, gammadion arranged, periodic nanostructures forming $\mathcal{F}$ / $\mathcal{P}$ unit cells. Main dimensions and material thickness’s are also depicted in the 3D model. 137

5.2 Composite SEM micrograph of ten different L-shaped, gammadion arranged, periodic nanostructures. The arrows indicate the varying $d$ distance between $\mathcal{F}$ unit cells, and fixed distance for $\mathcal{P}$ unit cell. Figure reproduced from Valev et al. [70]. 137

5.3 SHG microscopy measurement of the devices in Fig. 5.2. Incident laser excitation at 900 nm wavelength with right-CPL (top) and left-CPL (bottom). Red and yellow dashed squares highlight regions of strong SHG intensity corresponding to the centre of the gammadion structure with the same chirality as the incident CPL. Figure reproduced from Valev et al. [70]. 139
List of Figures

5.4 (a) SHG dichroism from $\mathcal{U}$ unit cell (yellow) and $\mathcal{P}$ unit cell (red), calculated through the ratio between the integrated intensity with left- and right-CPL for different gap distances $d$. (b) ratio of SHG intensity $I_L^{(2\omega)}/I_R^{(2\omega)}$ as a function of gap size $d$. The lines are guides to the eye. Figure reproduced from Valev et al. [70]. 140

5.5 Modelled reflection spectrum (at normal incidence) from a periodic arrangement of gammadion nanostructures with different gap sizes $d$. Left-CPL was modelled by setting the source to 45$^\circ$ polarization and a 90$^\circ$ phase difference. Numerical convergence was obtained with 12 diffraction orders. 142

5.6 Modelled spectral response for CPL excitation at 900 nm wavelength, as function of the gap size of the $\mathcal{U}$ nanostructure. Numerical convergence was obtained with 12 diffraction orders. 143

5.7 Spatial profile of the absolute electric field $|E|$ at the top surface Gold/Air interface of different gap size nanostructures. Top row shows electric field profile on $\mathcal{P}$ nanostructure excited with right-CPL, and bottom row shows $\mathcal{U}$ nanostructure excited with left-CPL. Numerical convergence was obtained with 26 diffraction orders. 144

5.8 Spatial profile of the absolute magnetic field $|H|$ at the top surface Gold/Air interface of different gap size nanostructures. Top row shows magnetic field profile on $\mathcal{P}$ nanostructure excited with right-CPL, and bottom row shows $\mathcal{U}$ nanostructure excited with left-CPL. Numerical convergence was obtained with 26 diffraction orders. 145

5.9 SAM flux at the gold-air interface of the nanostructures in Fig. 5.2, for normal incidence with right-CPL (top) and left-CPL (bottom) and 900 nm wavelength excitation. 148

5.10 Optical chirality enhancement $|\hat{C}(r)| > 1$ at the top gold/air interface for the (top) $\mathcal{P}$ nanostructure with right-CPL incidence and (bottom) $\mathcal{U}$ nanostructure with left-CPL incidence. Regions of enhanced optical chirality are located at the centre of the gammadion shaped structure. 151
5.11 Enhanced electric energy density $\hat{U}_e(r)$ at the top gold/air interface of the (top) $\mathcal{P}$ nanostructure with right-CPL and (bottom) $\mathcal{H}$ nanostructure with left-CPL. ............................................. 154
5.12 Dissymmetry factor $\hat{g}^*(r)$ at the top gold/air interface of the (top) $\mathcal{P}$ and (bottom) $\mathcal{H}$ nanostructures. .............................................................. 155
5.13 Spatial profile of the total SHG polarization $\mathbf{P}(2\omega, r)$ for (top) right-CPL and (bottom) right-CPL excitation on the (top) $\mathcal{P}$ and (bottom) $\mathcal{H}$ nanostructures with different gap sizes $d$. SHG profile was generated by finding the maximum polarization across the $z$-axis in the gold film. For comparison purposes between devices, the results were normalized with respect to a maximum polarization value. ... 157
5.14 SHG power from the (a) $\mathcal{H}$ nanostructure using left-CPL excitation and (b) $\mathcal{P}$ nanostructure with right-CPL. The colored data points correspond to (black) electric-dipole, (red) magnetic dipole and (green) electric quadrupole transitions. Yellow area corresponds to the total SHG power of all contributions. ............................................. 158
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Selection of EBL experiments using different process parameters.</td>
<td>47</td>
</tr>
<tr>
<td>3.2</td>
<td>Argon milling conditions and rates for different materials using a SVS 6000 sputtering system. Base pressure below $10^{-6}$mbar and sample rotation 58rpm.</td>
<td>52</td>
</tr>
<tr>
<td>3.3</td>
<td>Relevant material parameters for thermal evaporation. Extracted from the material deposition chart supplied by Kurt J. Lesker.</td>
<td>56</td>
</tr>
<tr>
<td>3.4</td>
<td>Process parameters for thermal evaporation of chromium, titatium and gold using a Edwards Auto 306 (resistive) and Edwards Auto 500 (electron beam) evaporators.</td>
<td>56</td>
</tr>
<tr>
<td>3.5</td>
<td>System specification for the Chameleon™ Ultra II laser used in this work.</td>
<td>68</td>
</tr>
<tr>
<td>3.6</td>
<td>Tensor components as function of different polarizer-analyzer configurations for normal incidence. The x and y components are relative to the plane of incidence of the sample.</td>
<td>77</td>
</tr>
<tr>
<td>4.1</td>
<td>MBE growth times, rates and final thickness for each sample and interface.</td>
<td>93</td>
</tr>
<tr>
<td>4.2</td>
<td>Average linear polarization density $\mathbf{P}(\omega, r)$, in units of Cm$^{-2}$, inside the ZnO film on the complementary samples. The values correspond to different LSP excitations.</td>
<td>104</td>
</tr>
<tr>
<td>4.3</td>
<td>Comparison of SHG and THG intensities at maximum average power (100mW) between complementary samples.</td>
<td>114</td>
</tr>
</tbody>
</table>
4.4 Comparison of SHG and THG intensities between complementary samples for two incident polarizations $\psi=[0^\circ,90^\circ]$ and an average power of 100 mW. Attenuation values calculated using the expression $\text{Att (dB)} = 10 \times \log_{10}(\max[I(\psi)]/\min[I(\psi)])$. 

4.5 Modelled transmitted SHG and THG power for two incident electric-field polarizations $\psi=[0^\circ,90^\circ]$ through (row 1) a Al$_2$O$_3$ substrate with an unstructured 75nm ZnO film, (row 2) through the transmissive sample and (row 3) through the reflective sample. The source was set to normal incidence $\theta=0^\circ$, electric-field amplitude $|E|=1$ V/µm, and wavelength $\lambda=1904$nm. Rows 4 and 5 show the enhancement owing to the LSP modes for the transmissive and reflective samples respectively. This is in relation to the measurements in row 1 i.e. ratio between sample and unstructured ZnO film. Rows 6 and 7 show the enhancement due to the change of the incident polarization $\psi$. Rows 8 and 9 show a comparison of the nonlinear response between complementary samples for polarizations with highest and lowest enhancement respectively.
Chapter 1

Introduction

Optics is a significant branch of physics and engineering which has been gaining huge momentum in the past few decades due to the large technological developments in the field of nanofabrication and also from our improved understanding of light-matter interactions. This has led us to significant technological milestones, such as the development of solar cells [1], lasers [2], optical fibres [3], CCD detectors [4], semiconductor lasers [5], ultrashort pulse lasers [6], photonic crystals [7, 8] and negative refractive index materials [9]. It also has been a driving source of innovation for a number of industries from display and lighting, to the manufacturing sector; from optical data communications to security and health care. The significance and potential of this area has led to huge capital investments in the United Kingdom; the Engineering and Physical Sciences Research Council has a portfolio of £250M\(^1\) to finance optics related topics, making 5.55% of the total portfolio (£4.5B in 2015).

The field of optical data communications is an area with large economic potential. There is a huge market demand for higher data transfer rates, over long distances and with reduced power consumption. Some of the short term technological efforts have been in optimizing data transfer rates by multiplexing different carrier signals over a single optical fibre using different light colors. Unfortunately, these systems

\(^1\)Grant values taken from https://www.epsrc.ac.uk/research/ourportfolio/vop/. The following research themes were considered: Light Matter Interaction and Optical Phenomena; Optical Communications; Optical Devices and Subsystems; Quantum Optics and Information; Photonics Materials and Metamaterials; and Optoelectronic Devices and Circuits
Chapter 1. Introduction

rely heavily on conversion links between electronic and optical signals, meaning that we are still limited by the response time of electronic circuits. In an ideal system, we would like to replace all electronic operations by an optical analogue. However, this is a very elaborate problem as it requires developing new components where light is allowed to interact with itself.

Intense light radiation when travelling through a material has an effect on the charge distribution and, consequently, on the optical properties of the material. This type of light-matter interaction is the sole subject of nonlinear optics. Shortly after the development of the first laser, the second harmonic generation process was experimentally demonstrated showing a controlled photon-to-photon interaction where two photons of frequency $\omega$ are converted to one photon of frequency $2\omega$. This discovery has opened the field of nonlinear optics and since then it has grown substantially with the development of a diverse range of tools and techniques such as tunable frequency radiation [10], all-optical switching [11], laser pulse shortening [12] and nonlinear optical spectroscopy [13].

Nonlinear optical effects are inherently weak and require a large amount of power to be excited, meaning that large optical systems are usually required, limiting their use in integrated optical systems. This is where the field of plasmonics comes into play. Plasmons are quantized electromagnetic oscillations of a free-electron cloud which are localized at the interface between a metal, and are very sensitive to changes in the dielectric and defects in the metal. Plasmons are responsible for the extraordinary optical transmission of light in subwavelength metal holes and beyond the diffraction limit [14]. This is now understood to be associated with large electromagnetic field enhancements, and has a direct application in nonlinear optics where effects have a power dependence on the fundamental field. Due to recent advancements in modelling and nanofabrication tools, we can now design and fabricate plasmonic nanostructures with interesting optical responses, and use this to miniaturize and perfect nonlinear systems. Some examples of nonlinear plasmonic devices include even-order nonlinear effects with no need for phase matching [15], bulk second-harmonic generation from materials with no birefringence [16], high-
harmonic generation [17], and optically switchable reflection and absorption [18].

Concepts from nonlinear plasmonics are not only useful in photonics but also find applications in diverse technological areas such as chemistry, biosensing and materials science. Nonlinear processes, such as second-harmonic generation (SHG), are a useful probe to study the surface chemistry of liquids, metals, oxides, polymers and semiconductors [19]. By monitoring the change in amplitude, polarization, phase or response times of surface SHG, it is possible to obtain information about absorption and surface coverage of a given species, or even the orientation of certain molecules. Surface SHG effects are intrinsically very weak in amplitude but can be enhanced significantly using plasmonic metasurfaces allowing a strong localization of the fields at the surface. Recent technological developments allow us to fine tune these effects and satisfy the requirements of a given application, providing a means of creating efficient, resourceful and compact sensing devices.

The main objective of the current work is to develop an understanding for the physics involved in the interaction of light with metallic nanostructures, and use this to develop novel devices with enhanced nonlinear optical effects. The core material of this thesis can be split in two research ideas: (1) investigation of surface SHG effects from chiral metasurfaces that introduce an extra enhancement through compression of the helicity of circularly polarized light (CPL); (2) integration of a nonlinear material on a metasurface with tunable linear optical response, and investigation of the effect of varying the plasmonic excitations on the generation of bulk nonlinear signals.

1.1 Outline

In Chap. 2, an overview of the main concepts and ideas in the field of plasmonics and nonlinear optics will be presented. This will be based on historical and current scientific work and will include discussion of some of the major developments in the field which are of relevance to this thesis.
Following this, Chap. 3 includes an introduction of the numerical methods used for studying the linear and nonlinear optical response of different metasurfaces. These were essential in the design of some of the plasmonic devices in this work allowing matching of experimental requirements, and in the understanding of experimental results through identification of dominant physical processes. An overview of the fabrication tools and measurement techniques will also be presented in this chapter.

Chapter 4 consists of a detailed report on the development of a plasmonic device with a tunable bulk SHG response from an isotropic noncentrosymmetric nonlinear material. Even-order nonlinear events can only occur in noncentrosymmetric materials and, typically, birefringence is also required for phase-matching to build a strong enough signal for real world applications. This in turn limits the potential of some materials which have good bulk nonlinear properties (e.g. GaAs and ZnO). Here, the effect of extraordinary optical transmission/reflection from complementary asymmetric cross metasurfaces is exploited by embedding them in a noncentrosymmetric material, allowing the generation of nonlinear signals without phase-matching. Specifically, it will be shown that control over the plasmonic modes, through change in polarization, fundamental wavelength or incident angle, allows a modification of the nonlinear response.

In Chap. 5 we look at surface SHG effects from chiral metasurfaces by excitation using circularly polarized light. Surface SHG effects are usually of lower magnitude in comparison to bulk ones, however, due to the field enhancement from plasmonic effects and the chiral nature of the metasurface, the nonlinear effects can be amplified and measured. Here, a numerical method was implemented to map regions of superchiral light with hotspots of strong SHG radiation, allowing to understand the physical processes involved in the dichroic response of the metasurfaces.

Finally, a summary of the main achievements of this thesis will be presented in Chap. 6 including key contributions to the field of nonlinear plasmonics and
discussion of some ideas to extend this work.
Chapter 2

Background

The optical properties of naturally occurring or chemically synthesized materials can be derived from a knowledge of their atomic and molecular structure. However, novel and unseen optical effects can be engineered by structuring and replicating the materials to dimensions below the wavelength of incident EM radiation. Meta-materials, where meta- comes from the Greek μετά- meaning beyond, is an emerging research area involved in the development of new artificial materials, and compromising of a range of disciplines including physics, chemistry, biology, materials science and engineering. Some of the early developments started at the end of the 19th century, as our understanding for electromagnetism reached new levels with James Clerk Maxwell’s work on “A Dynamic Theory of the Electromagnetic Field” [20], and led to some major scientific breakthroughs such as polarization rotation with chiral metamaterials [21], negative index materials [22, 23], and optical cloaking [24, 25]. Early works were limited to microwave radiation but with recent advancements in nanofabrication, we are now able to produce metamaterials with optical responses in the near-infrared (IR) and visible wavelength ranges, where there is great availability of source and detection technologies.

In this chapter, we will present two of the main branches of metamaterial research of relevance to this thesis, i.e. plasmonics and nonlinear metasurfaces. A description of the fundamental physical principles behind these topics will be provided, including highlighting and discussion of some of the major achievements in
2.1 Plasmonics

Plasmonic effects have been in use since the 4th century AD (at least) to produce a dichroic glass which can undergo a color change depending on the direction of light. The Lycurgus Cup, part of The British Museum collection, is a rare display of this Roman technology where gold and silver nanoparticles were added to the glass to exploit this unusual effect. However, it was not until more recent developments in classical electrodynamics [20], and the theoretical and experimental realization of plasma oscillations in metals [26], that we have gained a better understanding of the unusual optical effects from metallic nanoparticles. In this section, we will discuss the origin of the optical properties of metals, following a review of the fundamental principles for surface plasma oscillations in metals, including techniques for excitation of SPs and some of the major applications.

2.1.1 Dispersive permittivity of metals

The macroscopic EM response of a material can be derived from a knowledge of the relative values of permittivity $\varepsilon_r$ and permeability $\mu_r$ through calculation of the electric and/or magnetic polarizations. For most materials the magnetic polarization can be negligible for wavelengths below the near-IR range, thus the permeability is usually taken as $\mu_r = 1$. For dispersive media, such as metals, the permittivity is typically of the complex form, $\varepsilon_r = \varepsilon_{r1} + i\varepsilon_{r2}$, where the real part defines the degree of polarization induced by an external electric-field, and the imaginary part $\varepsilon_{r2}$ is a measure of the optical losses in the material. The sign of the real part of the permittivity gives an indication of the type of material, positive values for dielectric materials and negative for metals. In this section, we will discuss the origin of the dispersive permittivity for metals.

Metals are known for having a large number of free electrons, with densities
ranging from $n_e = 10^{28}$ to $10^{29}\text{m}^{-3}$, and when subject to an electric field, they can accelerate and undergo collisions with scattering times $\tau$ varying from $10^{-14}s$ to $10^{-13}s$. The oscillatory behaviour of the free electrons when subject to an external field $\mathbf{E}$ can be described by the equation of motion given by

$$\mathbf{F}(t) = m_e \frac{d^2 \mathbf{r}(t)}{dt^2} = -e\mathbf{E}(t) - m_e\gamma \frac{d\mathbf{r}(t)}{dt} \quad (2.1)$$

where $m_e$ is the electron mass, $\mathbf{r}(t)$ the electron position at time $t$, $e$ the elementary charge, and $\gamma = 1/\tau$ the damping frequency associated with the electron collisions due to the plasmon oscillation. The scattering time depends both on the material, and the shape and size of the plasmonic particle [27]. Eq. 2.1 can then be re-written as a function of polarization, a measure of charge separation in the material, using the relationship

$$\frac{d^2 \mathbf{P}(t)}{dt^2} = \frac{n_e e^2}{m_e} \mathbf{E}(t) - \gamma \frac{d\mathbf{P}(t)}{dt}. \quad (2.2)$$

If we consider the applied electric field to be monochromatic of angular frequency $\omega$, and with a harmonic time dependence $\mathbf{E}(t) = \mathbf{E}(\omega) \cos(\omega t) = \mathbf{E}(\omega)\Re\{\exp(i\omega t)\}$, then the polarization will be of the same form $\mathbf{P}(t) = \tilde{\mathbf{P}}(\omega)\Re\{\exp(i\omega t)\}$, where $\tilde{\mathbf{P}}(\omega)$ is a complex quantity which accounts for the phase difference with respect to the driving field $\mathbf{E}(\omega)$. Replacing these into Eq. 2.2 will lead us to

$$\tilde{\mathbf{P}}(\omega) = \varepsilon_0 \chi^{(1)}(\omega) \mathbf{E}(\omega) = \varepsilon_0 (\varepsilon_r(\omega) - 1) \mathbf{E}(\omega) = -\frac{n_e e^2}{m_e (\omega^2 - i\gamma \omega)} \mathbf{E}(\omega) \quad (2.3)$$

where $\varepsilon_0$ is the vacuum permittivity, $\chi^{(1)} = \varepsilon_r - 1$ the electric susceptibility and $\varepsilon_r(\omega)$ the complex dispersive permittivity. From this, we can then write an expression for the complex permittivity of a metal

$$\varepsilon_r(\omega) = 1 - \frac{\omega_p^2}{\omega^2 - i\gamma \omega} \quad (2.4)$$

where the angular frequency of the free electron gas, known as plasma frequency, is given by

$$\omega_p = \sqrt{\frac{n_e e^2}{m_e \varepsilon_0}}. \quad (2.5)$$

This result is also known as the Drude model [28], which attributes the optical losses in metals to intraband contributions from conduction-band electrons. Figure 2.1 shows the $\omega_p$ value for different metals which is typically within the ultra
violet (UV) range.

In Fig. 2.2, the relative permittivity obtained from the classical oscillator model is compared against experimental data for gold, which is a widely used plasmonic material due to its good optical/thermal properties and chemical stability [30]. Down to near-IR wavelengths there is a relatively good agreement between model and experimental data, however, in the visible and UV range the model breaks down due to interband transitions, i.e. as we reach the Fermi level of the material (for gold $E_F = 5.53 eV$ therefore $\lambda = 0.224 \mu m$) bound electrons are excited between the bands of the solid leading to optical transitions at different energies ($\hbar \omega_j$).

A formulation of the relative permittivity for interband transitions can be de-
Figure 2.2: Dispersive permittivity of gold. Individual contributions from the free-electrons (intraband, Drude model) and bound-electron (interband, Lorentz model) transitions are shown, including sum/total contributions [31]. Experimental data was taken from Palik [32].

...
bands can hold a large number of electrons (14 and 10 respectively) and as seen in Fig. 2.3, this translates into peaks in the total density of states confined to a narrow set of energies. These features correspond to allowed interband transitions, and explain the drop in reflectivity between 2eV and 5eV energies, as shown in Figure 2.3.(c). In the Lorentz model, these features are defined as separate oscillator terms in the equation. The interband contributions to the relative permittivity of gold is plotted in Fig. 2.2, and we can clearly see an increase in $\Im\{\varepsilon(\omega)\}$ as we reach the UV range, which is an indication of strong optical absorption.

A more complete description of the relative permittivity of metals is given by the Lorentz-Drude model, which includes both interband and intraband transitions, and it is expressed in the following form [31]

$$
\varepsilon_r(\omega) = 1 - \frac{f_0\omega_p^2}{\omega^2 - i\gamma_0\omega} + \sum_{j=1}^{k} \frac{f_j\omega_j^2}{\omega^2 - \omega_j^2 + i\gamma_j\omega}
$$

where $j$ is the oscillator number with frequency $\omega_j$, strength $f_j$ and inverse lifetime $\gamma_j$. The model parameters for different metals can be found in Rakic et. al. [31]. In Fig. 2.2, the full relative permittivity of gold is plotted using Eq. 2.7, showing a very good agreement with the experimental data.

### 2.1.2 Bulk plasmons

In equilibrium, a metal is characterized by having a fixed lattice of positive ions which are neutralized by an equal number of free electrons. If we imagine a situation where the free electrons are displaced by a distance $r'$, through a static electric-field ($\omega = 0$), a surface charge density is developed on either side of the metal with equal amplitude but opposite sign, i.e. $\sigma = \pm n_e e r' \text{ Cm}^{-2}$. In turn, the positive ions will exert a restoring force which opposes the electron displacement, through an electric field with value $E = n_e e r' / \varepsilon_0$. By turning off the static electric-field, the free electrons are allowed to oscillate against the positive background. This oscillation will occur as a longitudinal wave, meaning that the electric-field will take the same direction as the displacement force. The equation of motion for the free
Figure 2.3: (a) Band structure and (b) total density of states for bulk gold calculated using GPAW software [33]. The arrows highlight regions of interband transitions. The peaks in the density of states come from the 5d bands. (c) Reflectivity of gold obtained from the expression \( R = \left| \tilde{n} - 1 \right| \) where \( \tilde{n} \) is the complex refractive index converted from the relative permittivity of gold obtained from Rakic [31].

Electron gas in the absence of an external stimulus, can then be written as

\[
\mathbf{F} = m_e \frac{d^2 \mathbf{r}'}{dt^2} = -e \mathbf{E} = -\frac{n_e e^2}{\varepsilon_0} \mathbf{r}' \iff \frac{d^2 \mathbf{r}'}{dt^2} + \frac{n_e e^2}{\varepsilon_0 m_e} \mathbf{r}' = 0 \iff \frac{d^2 \mathbf{r}'}{dt^2} + \omega_p^2 \mathbf{r}' = 0
\]

which gives us the same result as in Eq. 2.5 for the angular frequency \( \omega_p \), that corresponds to the natural resonant frequency of the free electron gas as it is allowed to oscillate. The quantization of these oscillations is what we call bulk or volume plasmons, and they can be measured using electron energy loss spectroscopy (EELS). This technique consists of firing a beam of high-energy electrons \( (E_i \gg 1 \text{keV}) \) to-
wards a thin metallic film. The fast electrons excite bulk plasmons as they pass through the film, leading to a transmitted energy of $E_o = E_i - n\hbar\omega_p$ where $n$ is the total number of excitations. Both $E_o$ and $n$ are output parameters of the EELS technique, and so, the plasma frequency can be easily obtained from this relation.

Now we look into finding a dispersion relation for the free electron gas or bulk plasmon. For this we take the general Maxwell-Faraday and Maxwell-Ampère differential equations [34]

$$\nabla \times E(r, t) + \frac{\partial}{\partial t} B(r, t) = 0, \quad (2.8a)$$

$$\nabla \times H(r, t) - \frac{\partial}{\partial t} D(r, t) = J(r, t) \quad (2.8b)$$

where $B$ is the magnetic flux density, $H$ the magnetic field strength, $D$ the electric displacement, and $J$ the current density, all functions of the position vector $r$ and time $t$. To combine both equations, a new curl operator and the known relations $B = \mu H$ and $D = \varepsilon E$ are introduced in Eq. 2.8a, leading to

$$\nabla \times \nabla \times E(r, t) + \mu \frac{\partial}{\partial t} \nabla \times H(r, t) = 0, \quad (2.9)$$

$$\nabla \times H(r, t) - \varepsilon \frac{\partial}{\partial t} E(r, t) = J(r, t). \quad (2.10)$$

The parameters $\varepsilon$ and $\mu$ correspond respectively to the permittivity and permeability of the medium. These can also be expressed in relative terms to the vacuum permittivity and permeability constants ($\varepsilon_0$ and $\mu_0$), using the relationships $\varepsilon = \varepsilon_0 \varepsilon_r$ and $\mu = \mu_0 \mu_r$. Replacing $\nabla \times H$ from Eq. 2.10 into Eq. 2.9, will lead to

$$\nabla \times \nabla \times E(r, t) + \mu \frac{\partial}{\partial t} J(r, t) + \varepsilon \mu \frac{\delta^2}{\delta t^2} E(r, t) = 0. \quad (2.11)$$

We want to find a solution for a travelling-wave with no external stimulus, and so a simplification of Eq. 2.11 can be made by setting the source term $J$ to zero. Also, replacing the curl of a curl of the $E$ vector with the relation $\nabla \times \nabla \times E = \nabla(\nabla \cdot E) - \nabla^2 E$, will lead to

$$\nabla(\nabla \cdot E(r, t)) - \nabla^2 E(r, t) + \varepsilon \mu \frac{\delta^2}{\delta t^2} E(r, t) = 0. \quad (2.12)$$

Considering the plane wave definition of the electric-field $E(r, t) = E(k, \omega) \exp(i(k \cdot r + \omega t))$, where $k = 2\pi/\lambda \hat{e}_\perp$ is the wave vector, $\omega$ the angular frequency, and $\hat{e}_\perp$ a
normed vector space which can take any angular deviation relative to the coordinate axis. Equation 2.12 can then be re-written in the Fourier domain by replacing $\partial/\partial t \rightarrow i\omega$ and $\nabla \rightarrow ik$, leading to

$$k(k \cdot E(k, \omega)) - k^2 E(k, \omega) + \varepsilon(\omega)\mu \omega^2 E(k, \omega) = 0.$$ 

For the case of a transverse wave, where oscillation of the electric-field is perpendicular to the direction of propagation of the wave vector, the simplification $k \cdot E = 0$ can be made, leading to

$$ck = \sqrt{\varepsilon_r(\omega)\mu_r \omega} \quad (2.13)$$

where $c = 1/\sqrt{\varepsilon_0\mu_0}$ is the speed of light. For the case of free-propagating light in vacuum, $\varepsilon_r = \mu_r = 1$, and the dispersion relation has a linear relationship with angular frequency, as illustrated in Fig. 2.4 by the $ck = \omega$ function.

Figure 2.4: Dispersion curve for the free electron gas and free propagating light in vacuum.

Considering the regime where the angular frequency $\omega$ of an EM-wave propagating in a metal is very close to the plasma frequency $\omega_p$, with values in the order of $10^{16}$ rad/s as shown in Fig. 2.1, and given that the damping rates in metals are in the order of $10^{12}$ to $10^{13}$ Hz [31], we can make a simplification of the Drude model.
in Eq. 2.4 by neglecting damping effects

\[ \varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2}. \]  

(2.14)

For the simplified case where no interband (bound-electron) transitions are allowed, replacing Eq. 2.14 into Eq. 2.13 and considering a non-magnetic metal \( \mu_r = 1 \), we can obtain an expression for the dispersion relation of the free electron plasma

\[ \omega = \sqrt{c^2k^2 + \omega_p^2}. \]  

(2.15)

Figure 2.4 shows a plot of the dispersion relation. For frequencies in the region \( 0 < \omega < \omega_p \), waves incident to the metal will not propagate and will be totally reflected, which is why metals are good reflectors in the visible to near-IR range. Propagation or transmission of light as a transverse wave is only possible for frequencies above the plasma frequency \( \omega > \omega_p \) where the metal behaves as a transparent dielectric. However, due to differences in the wave vector, bulk plasmons cannot be excited with free propagating light.

### 2.1.3 Surface plasmons

Surface plasmons are associated with oscillation of free electrons at the interface between a metal and a dielectric. In contrast to bulk plasmons, SPs have both longitudinal and transverse components which makes them extremely useful as they can interact with free propagating light (of transverse nature). The coupling between an SP and a photon leads to what we call an surface plasmon polariton (SPP) which has a greater momentum than a photon of the same wavelength. The electromagnetic field associated with the SPPs is bound and propagates along the surface of the metal until all its energy is dissipated as heat in the material.

### Case of p-polarized wave on a metal-dielectric interface

For the case presented in Fig. 2.5 of an SP propagating in the x-direction as a p-polarized (TM) plane wave (i.e. \( E_y = H_x = H_z = 0 \)), localized to the interface between a metal and dielectric \( (z = 0) \), and with electromagnetic fields decaying
Figure 2.5: Field distribution of an SP propagating as a p-polarized or TM wave, i.e. no magnetic field allowed in the direction of propagation.

exponentially along the z-axis in the half-spaces, we can write the following wave equations for each side of the interface

\[ \mathbf{E}^{d,m} = \left( E_x^{d,m} i \pm E_z^{d,m} k \right) \exp \left( i(k_x x - \omega t) \right) \exp \left( i k_{d,m} z \right) \tag{2.16a} \]
\[ \mathbf{H}^{d,m} = H_y^{d,m} \exp \left( i(k_x x - \omega t) \right) \exp \left( i k_{d,m} z \right) \tag{2.16b} \]

and define the following boundary conditions at \( z = 0 \)

\[ \varepsilon^d E_z^d = \varepsilon^m E_z^m, \tag{2.17a} \]
\[ E_x^d = E_x^m = E_x, \tag{2.17b} \]
\[ H_y^d = H_y^m = H_y, \tag{2.17c} \]
\[ k_x^d = k_x^m = k_x. \tag{2.17d} \]

Superscripts \( \{m, d\} \) reference the medium (metal or dielectric), the term \( \exp \left( i(k_x x - \omega t) \right) \) defines a monochromatic plane wave with \( k \) vector propagating in the x-axis and frequency \( \omega \), and \( \exp \left( i k_{d,m} z \right) \), is a term that defines the exponential decay of the electromagnetic fields in the z-axis.

In the absence of an external source (\( \mathbf{J} = 0 \)), if we replace Eq. 2.16 in Maxwell-Ampère equation (Eq. 2.10) and take into account the boundary conditions shown
Chapter 2. Background

In Eq. 2.17, we get the following pair of equations for each half-space

\begin{align*}
\text{dielectric: } & \quad + k_z^d H_y = \omega \varepsilon_0 \varepsilon_r^d E_x, \quad (2.18a) \\
& \quad k_x H_y = -\omega \varepsilon_0 \varepsilon_r^d E_z, \quad (2.18b) \\
\text{metal: } & \quad - k_z^m H_y = \omega \varepsilon_0 \varepsilon_r^m E_x, \quad (2.18c) \\
& \quad k_x H_y = -\omega \varepsilon_0 \varepsilon_r^m E_z^m \quad (2.18d)
\end{align*}

Then, from Eq. 2.18a and Eq. 2.18c we can write the following relation

\[ \frac{k_z^d}{k_z^m} = -\frac{\varepsilon_r^d}{\varepsilon_r^m}, \quad (2.19) \]

Surface confinement requires that \( \text{Re}\{k_z^d\} > 0 \) and \( \text{Re}\{k_z^m\} > 0 \) and since for most dielectrics \( \varepsilon_r^d > 0 \), the only way to satisfy Eq. 2.19 is by having \( \text{Re}\{\varepsilon_r^m\} < 0 \) which is a defining condition for most metals. This condition shows that surface plasmons can only occur in the interface between a metal and a dielectric, and are strongly dependent on the relative permittivity of both materials.

Continuing our derivation, conservation of the wave vector requires that

\begin{align*}
\text{dielectric } k_z^d = & \quad \varepsilon_r^d \frac{\omega^2}{c^2} = k_x^d \omega^2 + k_y^d, \quad (2.20a) \\
\text{metal } k_z^m = & \quad \varepsilon_r^m \frac{\omega^2}{c^2} = k_x^m \omega^2 + k_y^m \quad (2.20b)
\end{align*}

And so, combining Eq. 2.19 and Eq. 2.20 will lead us to

\begin{align*}
\frac{k_z^d}{k_z^m} = & \quad \varepsilon_r \frac{\varepsilon^d}{\varepsilon_r^d + \varepsilon_r^m}, \quad (2.21a) \\
\frac{k_z^m}{k_z^d} = & \quad \varepsilon_r \frac{\varepsilon^m}{\varepsilon_r^d + \varepsilon_r^m}, \quad (2.21b) \\
k_z^x = & \quad \varepsilon_r \frac{\varepsilon^d \varepsilon^m}{\varepsilon_r^d + \varepsilon_r^m} \quad (2.21c)
\end{align*}

As to be seen in the following sections, Eq. 2.21c can be used to calculate the dispersion relation and propagation length of a SP, which is essential for realising alternative coupling mechanisms, and in the selection of plasmonic materials based on wavelength requirements. Furthermore, Eq. 2.21a and Eq. 2.21b can be used to predict the decay length of a SP using different dielectric and metallic interfaces. An awareness of the decay length of SP is of particular interest for determining the
maximum useful thickness of a nonlinear dielectric material.

**Dispersion relation of surface plasmons**

In Fig. 2.6, we show the dispersion relation for an SP based on Eq. 2.21c. Some assumptions were made for this case, including a very small damping of the free-electron gas ($\gamma = 0$), a vacuum electric permittivity $\varepsilon_{r}^d = 1$, and a simplified Drude model equation for the relative permittivity of metals (Eq. 2.14). The dispersion relation shows that for frequencies below the plasma frequency ($\omega < \omega_p$), a p-polarized EM wave can propagate along the metal-dielectric interface up to $\omega_{sp} = \omega_p/\sqrt{1 + \varepsilon_{r}^d}$. Above the plasma frequency ($\omega > \omega_p$), we have seen in Sec. 2.1.2 that the metal becomes transparent, allowing radiation across the film. In the frequency interval $\omega_{sp} < \omega < \omega_p$, the wave vector $k$ is purely imaginary, not allowing EM propagation. For real metals this is not the case since interband transitions play a role in frequencies near $\omega_p$, and so the wave vector gains a real finite value.

Another observation we can make from Fig. 2.6 is that the wave vector of an SP is larger than of free-propagating light, not allowing a direct coupling between photons and SPs. Phase matching of the waves is only possible by reducing the speed of free propagating light, through introduction of an optical prism or grating. This will be discussed in more detail in Sec. 2.1.3. Also to highlight in Fig. 2.6 is the change of the wave vectors proportional to the dielectric permittivity $\varepsilon_{r}^d$, showing that the SP response is very much dependent on the optical properties of the dielectric medium. This is one of the key features used in sensing for plasmonic applications, as will be discuss further in Sec. 2.1.6.

**Surface plasmon decay and propagation length**

Coming back to our solution of an s-polarized wave propagating in a metal-dielectric interface, it is possible to calculate the z-decay length (skin depth) of the SP on either side of the interface through the reciprocal relation $L_{z}^{d,m} = |\Im[k_{z}^{d,m}](\omega)|^{-1}$.
Figure 2.6: Dispersion curve of an SP, bulk plasmon and free propagating light using two different dielectric materials, vacuum \((\varepsilon^d_r = 1)\) and \(\text{CaF}_2\) \((\varepsilon^d_r = 2)\). The Drude model for ideal metals (no losses, \(\gamma = 0\)) was considered in the description of the dispersive relative permittivity \(\varepsilon^m_r(\omega)\) (Eq. 2.14).

Similarly, the propagation length of the SPs, limited by the optical losses in the metal \((\Im\varepsilon^m_r(\omega) \neq 0)\), can be obtained from \(L_z(\omega) = |2\Im[k_x(\omega)]|^{-1}\).

Figure 2.7 shows the typical propagation and decay length values for SPs propagating in a gold-vacuum interface. The decay length inside the metal \((L_z^m)\) does show a substantial variation of magnitude across a range of wavelengths. However for the propagation length \((L_z)\) and decay length in the dielectric \((L_z^d)\), we see a substantial increase in values as we move away from the SP frequency \((\omega_{sp})\), suggesting a strong field confinement at smaller wavelengths.

\textbf{s-polarized wave on a metal-dielectric interface}

Similarly to Sec. 2.1.3, we can also consider the case of a s-polarized (TE) surface wave propagating in the x-direction and with a field decay in the z-axis. The wave
Figure 2.7: Calculation of the SP propagation length $L_x$ and $z$-decay length inside a dielectric $L_d^z$ and a metal $L_m^z$. The mediums on each half of the interface are vacuum ($\varepsilon_r^d = 1$) and gold. The Lorentz-Drude model in Eq. 2.7 was used to calculate the relative permittivity $\varepsilon_r^m$ of gold.

Equations for this polarization are as follows

$$E_{d,m} = E_{y,d,m}^d i \exp(i(k_x x - \omega t)) \exp(ik_z^d z),$$

$$H_{d,m} = (H_{x,LM}^d \pm H_{z,LM}^d) \exp(i(k_x x - \omega t)) \exp(ik_z^d z).$$

Replacing these into the Maxwell-Faraday equation (Eq. 2.9), and finding a solution considering the boundary conditions $H_x = H_{d,m}^d = H_{m}^m$, $H_z = \mu_d^d H_{x}^d = \mu_m^m H_{x}^m$ ($\mu^d = \mu^m = 1$ for non-magnetic materials), $E_y = E_{y,d,m}^d = E_{m}^m$ and $k_x = k_{x,d}^d = k_{x,m}^m$ at $z = 0$ will lead to

- dielectric: $+ k_{d,z}^d + k_x = 0,$
- metal: $- k_{m,z}^m + k_x = 0.$

Conservation of the wave vector requires that $k_{d,z}^d + k_{m,z}^m = 0$, but we are not able to satisfy this condition since it is required that $\Re[k_{d,z}^d] > 0$ and $\Re[k_{m,z}^m] > 0$, meaning that surface EM waves cannot propagate with s-polarization. Alternatively, excitation of p- or s-polarized waves can be achieved by using a dielectric material with Kerr-type nonlinearities [35, 36]. This is characterized for having an intensity-dependent
relative permittivity given by \( \varepsilon_d = \varepsilon_{d,0} + \alpha |E(\omega)|^2 \), where \( \varepsilon_{d,0} \) and \( \alpha \) are the weak-field and nonlinear permittivities, and \( I \propto |E|^2 \) the electric field amplitude proportional to the incident intensity.

**Techniques for surface plasmon polariton coupling**

The dispersion relation in Fig. 2.6 shows that a surface EM wave has a wave vector larger than that of free propagating light, and so, momentum conservation does not allow a direct coupling of SPs with photons. However, some techniques are available which enhance the wave vector of the incident photons, and thus meet the phase matching required for SPP excitations.

Early experiments involved the use of EELS for the excitation and detection of SPP, including mapping of the full wave dispersion [37]. This technique uses high energy electrons to excite surface modes in thin metallic films, and measures their energy and momentum loss through detection of the transmitted electrons. More sophisticated systems can now resolve the field profiles of individual plasmonic modes from single metallic nanoparticles [38].

A simpler technique uses an optical prism to modify the wave vector of the incident light [39, 40]. Total internal reflection is achieved inside the prism which induces evanescent waves with a wave vector \( k_p = \frac{\sqrt{\varepsilon_p \omega}}{c} \sin \theta_i \), where \( \varepsilon_p \) is the relative permittivity of the prism and \( \theta_i \) the incidence angle. There are essentially two types of configurations using an optical prism that allow to exploit a limited range of SPP modes (Fig. 2.8.(a)). The Otto configuration [39] involves having the prism in close proximity to a metallic film, with the gap filled with a lower refractive index material (\( \varepsilon_g < \varepsilon_p \)) and a distance within the decay length of the evanescent wave and the SPP (typically a few hundreds of nanometers). For \( \sqrt{\varepsilon_p} \sin \theta_i > \sqrt{\varepsilon_g} \) (or \( \arcsin(\sqrt{\varepsilon_g/\varepsilon_p}) \leq \theta_i \leq \pi/2 \)), the evanescent wave will propagate across the gap and excite an SPP with wave vector \( k_{\text{SPP}}(\omega) = \varepsilon_p \omega/c \sin^{-1} \theta_i \). Alternatively, there is the Kretschmann and Raether configuration [40] where a thin metal film is directly evaporated on the optical prism with a thickness smaller than the decay
length of the evanescent wave in the metal. In this case, the evanescent wave propagates across the metal from the optical prism, using again a \( \theta_i \) angle larger than the critical angle for total internal reflection. It then excites an SPP mode in the metal-dielectric interface. The reason to use a prism rather than a planar waveguide of the same dielectric material is to avoid total internal reflection within the slab. Once an SPP mode is excited, inversely through the same process, the SPP can generate a plane wave that comes out on the other side of the prism. In this situation, adding to the intrinsic losses in the metal described by \( \Im[\varepsilon_r(\omega)] \), there will be radiative damping due to the energy conversion between the SPP and the plane wave, and so the SPP is detected as a minimum in the reflected light.

![Figure 2.8: Dispersion curves for SPP using the (a) prism and (b) grating coupling methods.](image)

For the optical prism case, phase matching occurs for incident angles \( \arcsin(\sqrt{\varepsilon_g/\varepsilon_p}) \leq \theta_i \leq \pi/2 \) where evanescent waves and SPP wave vectors are equal.

Another coupling method involves using scattering defects on a metallic surface. Earlier work used periodic metallic gratings to decompose the incident light into components with varying wave vectors [41]. Figure 2.8.(b) shows the folding dispersion curves for the scattered SPP fields as result of the periodic modulation introduce by the grating, i.e. \( k_\parallel = k_{\text{SPP}} \pm nk_g \) where \( n \in \mathbb{Z} \) and \( k_g = 2\pi/P \) is the grating wave vector of period \( P \). The complex band structure formed by the grating leads to the excitation of SPP when \( |k_\parallel| \leq \omega/c \). Alternative structures
based on scattering defect coupling include periodic nano-hole arrays [42] and zig-zag gratings [43].

2.1.4 Localized surface plasmons

In this section we will be looking into a different type of plasmons known as LSPs. In contrast to SPPs which are propagating EM waves, LSPs are non-propagating excitations of the free electron gas, confined to the vicinity of isolated metallic structures, with a length scale comparable to the size of the object. LSPs have a natural resonant frequency which is not only dependent on the dielectric properties of the metal and surrounding medium, but also on the shape and size of the structure [44]. At resonance, the LSPs allows the confinement of light into small volumes and are associated with the phenomenon of extraordinary optical transmission (EOT) through subwavelength holes in thin metallic films [45]. Another relevant feature of LSPs is that coupling with free propagating light can be achieved without the need of phase matching between wave vectors, due to the translational asymmetry of the fields in metallic structures [46].

Consider the case of an EM-wave propagating towards a metallic sphere with dispersive permittivity \( \varepsilon_m(\omega) \), radius smaller than the wavelength of light \( r \ll \lambda \), and surrounded by a dielectric medium with permittivity \( \varepsilon_d \) (inset in Fig. 2.9). The EM field can lead to the oscillation of conduction electrons in the metal, creating a displacement of charges around opposite sides of the sphere. Coulomb interactions between different charges give rise to a restoring force of the conduction electrons, with a natural resonant frequency \( \omega_{\text{LSP}} \). Efficient energy coupling to the LSP can be achieved when the frequency of light matches the resonance \( \omega_{\text{LSP}} \). The Rayleigh approximation is a common method to study this problem, limited however to metallic spheres of size \( d \leq \lambda/20 \). A solution to the polarizability of a metallic sphere using this approximation has been derived in Jackson [34] and is given by

\[
\alpha(\omega) = 4\pi r^3 \frac{\varepsilon_m(\omega) - \varepsilon_d}{\varepsilon_m(\omega) + 2\varepsilon_d}. \tag{2.22}
\]

An interesting event occurs when the denominator \( \varepsilon_m(\omega) + \varepsilon_d \) is at its minimum,
Chapter 2. Background

Figure 2.9: Plot of the absolute and phase values of the polarizability of a gold sphere embedded in vacuum (solid line, \( \varepsilon_d = 1 \)) and water (dashed line, \( \varepsilon_d = 1.8 \)), calculated using Eq. 2.22. Dispersive permittivity of gold was obtained from the Lorentz-Drude model in Eq. 2.7 using the parameters in Rakic et al. [31]. Inset shows a sketch of the gold sphere subject to an electric-field.

Corresponding to a resonant enhancement of the LSP. This condition is achieved for \( \Re[\varepsilon_m(\omega)] = -2\varepsilon_d \) which requires a material with negative permittivity. Figure 2.9 shows the absolute and phase values of the polarizability for a gold sphere of radius 25 nm with resonances at \( \lambda = 526 \text{ nm} \) and \( \lambda = 543 \text{ nm} \), showing that as we increase the relative permittivity of the dielectric medium \( \varepsilon_d \), the electric field within the dielectric increases to counterbalance the local field in the metallic sphere, resulting in a red-shift of the LSP resonance towards higher wavelengths (lower frequencies).

Under plane-wave illumination, the polarizability of the sphere \( \alpha(\omega) \) leads to an oscillating dipole moment at the centre of the sphere with a time dependence \( p(r, t) = \varepsilon_0 \varepsilon_m \alpha(\omega) E_0 \exp(-i\omega t) \). This changes spatially with the incident EM-field polarization [47]. However, due to the symmetry of the metallic sphere the far-field response is independent of the incident polarization. The oscillating dipole is responsible for scattering effects in the sphere, which can be quantified using the
Chapter 2. Background

Figure 2.10: Scattering cross section calculated using Eq. 2.23 for a 50 nm gold sphere in a water based solution ($\varepsilon_d = 1.8$). The color of the dispersed gold sphere solution is determined by the difference in absorbed (not shown) and scattered cross sections. The scattering cross section equation from Bohren and Huffman [48], written as

$$C_{\text{sca}} = \frac{k^4}{6\pi} |\alpha(\omega)|^2. \quad (2.23)$$

This quantity provides a measure of how strongly light of a specific frequency is reflected from a metallic nanoparticle, showing increased scattering effects for particles with a larger surface area ($C_{\text{sca}} \propto r^6$). It is one of the parameters that defines the color of dispersed solutions of metallic nanospheres (Fig 2.10). Similar analysis can be made to the complementary geometry (subwavelength holes in a thin metallic film) in which case the polarizability is given also by Eq. 2.22 through substitution of $\varepsilon_d \rightarrow \varepsilon_m(\omega)$ and $\varepsilon_m(\omega) \rightarrow \varepsilon_d$, thus leading to the resonant condition $\Re[\varepsilon_m(\omega)] = -\varepsilon_d/2$.

2.1.5 Plasmonic materials

As discussed in Sec. 2.1.3 and 2.1.4, a plasmonic material requires having a negative real permittivity ($\Re[\varepsilon_r] < 0$) which is a characteristic of some bulk metals. A
limiting factor is typically the high loss given by the imaginary part of the relative permittivity \( \Im[\varepsilon_r] \), which affects the quality factor of the plasmonic resonance and the wavelength operation range of the material. Comparison studies with an perfect electric conductor (PEC) (no optical losses thus \( \Im[\varepsilon_r] = 0 \)), show however stronger plasmonic effects from real metals with low optical losses. This is because of the allowed penetration of the EM-field in real metals, which enlarges the effective area of the material, and therefore increases the transmission of the signal [49]. As a rule of thumb, aluminium is typically the preferred material for UV applications [50], silver for visible wavelengths [49], and gold for near- and mid-IR frequencies. Different types of alloys, intermetallics, silicides, high pressure materials, liquid metals and amorphous materials have also been investigated for plasmonic applications [51]. The alloys, silicides and some intermetallic materials have been found to have partially occupied \( d \) states which brings the probability of interband transitions to the lower frequency range limiting the use of these materials in plasmonics. Also some doped and amorphous metal alloys show higher optical absorption due to increased scattering effects. Intermetallics with a smaller number of atoms in the unit cell and a simpler band structure were found to have a performance comparable to silver or gold, but at the cost of more complex deposition techniques.

Alternatively to bulk materials, plasmonic effects have been studied in two-dimensional electron gas (2DEG) interfaces such as semiconductor heterojunctions [52, 53, 54, 55] and 2-dimensional electronic systems (e.g. graphene [56, 57, 58, 59, 60, 61, 62, 63, 64, 65]). These are also characterized for having a large electron density (in the order of \( n_e=10^{18} \text{ m}^{-2} \)) which is subject to Coulomb restoring forces from the positive charged background, and therefore can support plasmonic waves. One of the key features of 2DEG materials is that the dispersion relation does not exhibit an asymptotic behaviour as bulk metals, and so plasmonic excitations can occur at a wide range of frequencies. Another advantage of 2DEG materials is the possibility of modifying the dispersion of SPPs through control of Coloumb interactions using a gate voltage.
2.1.6 Plasmonic phenomena and applications

The mechanism by which light can couple with the conduction electrons in metallic films can lead to interesting and unusual optical effects. In this section, the most relevant effects for the thesis will be discussed, with references to relevant literature and potential applications.

Chiral metasurfaces

As discussed in Sec. 2.1.4, LSPs can be excited with free-propagating light in sub-wavelength metallic particles and planar geometries (or metasurfaces), offering great potential in the manipulation of the properties of light. One area of extensive research is in chiral metasurfaces, composed of metallic nanoparticles that lack mirror image symmetry allowing them to interact differently with left- or right-circularly polarized light (also of chiral nature). Chiral optical effects such as optical rotation and circular dichroism, have been extensively studied in crystals and organic liquids since the beginning of the XIX century [66, 67]. However, it was not until recently that a planar metasurface was demonstrated to change the polarization state of CPL [68]. Similarly to natural chiral optical materials, this new metasurface allows mixing of the electrical and magnetic components of light through excitation of magnetic and electric dipoles in the material respectively. With metamaterials we can however manipulate the EM interaction through excitation of LSPs leading to enhancements of the EM response, and geometrical confinement allowing efficient coupling of the EM-fields. One relevant application is a metamaterial that can compress the helical pitch of CPL to the size of small molecules, allowing detection of their chiral optical response [69]. This effect is strongly associated with enhanced optical near-field intensities, and so, chiral metamaterials offer great potential in the field of nonlinear optics. Some examples include metamaterials with enhanced nonlinear optical rotation and circular dichroism [70, 71] which form part of Chap. 5 of this thesis.
Extraordinary optical transmission

The effect of extraordinary optical transmission (EOT) through subwavelength periodic holes in a thin metallic film [14] has triggered a huge interest in the scientific community since it was first discovered. Early work on the diffraction of light through single holes on an infinitely thin metallic film, modelled transmission as a process dominated by evanescent fields, whose magnitude is proportional to \( T \propto \left(\frac{d}{\lambda}\right)^4 \). This implies that for holes of diameter smaller than the wavelength of light \((d \ll \lambda)\), the transmission is considered negligible [72]. For real situations where the metallic film has a finite thickness \(t\), waveguide modes can exist in the cylindrical hole allowing a second transmission mechanism up to a cut-off wavelength \(\lambda_c > 2t\), still larger than the diameter of the aperture [73].

Further enhancements in transmission below \(\lambda_c\), can be achieved using a periodic arrangement of subwavelength circular holes in a metallic film [14, 74]. The hole lattice acts as a mechanism for excitation of SPPs, similarly to the optical gratings discussed in Sec. 2.1.3, and each hole allows tunnelling of the SPP from the incident to the opposite surface of the film, where the EM-field is then scattered away to the far field preserving the original characteristics of the incident light. In this arrangement, the periodicity of the array and the aspect ratio of the cylindrical holes \((t/d)\) are the major parameters influencing respectively the spectral position and shape of the SPP resonances [14, 75, 76]. Since each aperture channels more light than is incident on its area, the observed effect was coined extraordinary optical transmission. Similar effects were also observed with single hole apertures surrounded by periodic gratings [77], including control over emission on the exit side of the aperture by patterning both sides of the film [78].

The most obvious application for EOT is wavelength filtering, given the simplicity to produce devices with a desirable optical response. However, the EM-field confinement from an SPP excitation and its sensitivity to environment changes, are features of relevance for sensing applications. Some examples include excitation of vibrational molecular modes [79, 80], enhancement of fluorescence effects in spec-
troscopic techniques [81, 82], monitoring of surface molecular binding [83], and in nonlinear optics [16].

**Tunable plasmonic nanostructures**

Further studies on EOT using metallic apertures with an in-plane asymmetry, such as a rectangular/slit hole, revealed that transmission can be sensitive to the polarization of the incident light. From measuring light transmission through single apertures, where no propagating SP can exist, it was found that the optical response was characterized by the excitation of non-propagating LSP resonances around the aperture [84]. Further EELS measurements on individual holes confirmed that the dipole formed by the LSP is polarization sensitive [84]. The transmission peak from LSP excitation is essentially associated with Fabry-Perot interferences in the inside walls of the apertures, which occur near the cut-off wavelength of the waveguide mode. In comparison to PEC materials, the evanescent nature of the SPPs in real metals leads to a larger effective diameter of the aperture, resulting in a redshift of the cut-off wavelength [85].

The polarization dependent coupling of LSP modes has been observed in geometries such as elliptical [86, 87], rectangular [84, 88, 89] and asymmetric cross shaped apertures [90, 91, 92, 93, 94, 95, 80]. Some of the applications for this functionality include optical switching [87], quarter wave plates [96], tunable molecular vibrational mode interactions [80], dynamic color pixels [97], and tunable nonlinear optical response.

Alternatively to using the polarization of light as a control mechanism for the plasmonics modes, 2DEG materials offer great potential in the area of active tunable plasmonics [52, 53, 54, 65]. The charge density of the 2DEG can be adjusted by applying a gate voltage to the 2DEG, changing the relative permittivity of the material and so affecting the plasmonic response.
Babinet’s principle in complementary metasurfaces

An interesting concept in optics is Babinet’s principle which stipulates that the sum of diffracted waves from an opaque object and its complementary aperture is equal to the undistorted incident wave [98]. Two objects are said to be complementary to each other when an opaque region of one object corresponds to a transparent region of the other (and vice-versa, Fig. 2.11).

Figure 2.11: Incident and transmitted EM-fields through complementary objects.

The EM form of Babinet’s principle was first demonstrated by Booker [99] by considering two objects made on an infinitely thin PEC. As illustrated in Fig. 2.11, the original object consists of an aperture in the PEC film, while the complementary is an opaque obstacle both positioned at \( z = 0 \). A solution to Babinet’s principle requires that the sources behind the original and complementary objects \( (z < 0) \) produce a incident EM field that is related by \( \mathbf{E}_i^c = u \mathbf{H}_i^c \) and \( \mathbf{H}_i^c = -\mathbf{E}_i^o / u \), where \( u \) is in units of \( \Omega \), allowing an unit relationship between the electric \( \mathbf{E} \) and magnetic \( \mathbf{H} \) fields. The incident fields will then induce a surface-charge density and surface-current density that lead to scattered fields \( \mathbf{E}^s \) and \( \mathbf{H}^s \). In the \( z < 0 \) region, \( (\mathbf{E}^s, \mathbf{H}^s) \) are known as reflected fields, while for \( z > 0 \) these are diffracted fields. With exception to regions inside the PEC film, where the field is not allowed to penetrate, the total EM is a superposition of the incident and scattered fields, expressed by \( \mathbf{E}^t = \mathbf{E}^i + \mathbf{E}^s \) and \( \mathbf{H}^t = \mathbf{H}^i + \mathbf{H}^s \). A consequence to Babinet’s principle leads to a solution that relates the total EM fields between complementary samples at \( z > 0 \)
with the incident field behind the original source [34, 100, 101], these being

\[
E_{o}^{t} - uH_{o}^{t} = E_{o}^{i}, \tag{2.24}
\]

\[
H_{o}^{i} + E_{o}^{t}/c = H_{o}^{i}, \tag{2.25}
\]

where also at the region \( z > 0 \), the total fields of the original object \((E_{o}^{t}, H_{o}^{t})\) and the scattered fields of the complementary object \((E_{c}^{s}, H_{c}^{s})\), are related by \( E_{o}^{t} = uH_{c}^{s} \) and \( H_{o}^{i} = -E_{c}^{t}/c \).

Interestingly Babinet’s principle also holds for subwavelength plasmonic structures which lack perfect conductivity and ideal boundary conditions [102, 103, 104]. It was shown that complementary structures have complementary optical responses (transmission interchanges with reflection), field profiles and effective permittivity relationships when excited with orthogonal plane wave polarizations. Application of Babinet’s principle is of relevance in areas requiring indirect mapping of the magnetic fields, for which current probe technology is still very limited (in particular in the THz regime) [105].

2.2 Nonlinear optics

Light that travels through a material can interact with the electrons in the medium and induce a change in the distribution of charges (polarizability). At low intensity values, the polarizability \( P \) has a linear relation with the electric field \( E \) and oscillates with identical frequency \( \omega \). This is described by the relation \( P(r, \omega) = \epsilon_0 \chi_{ij}^{(1)} E(r, \omega) \) where \( r \) is a position vector, \( \epsilon_0 \) is the vacuum permittivity and \( \chi_{ij}^{(1)} \) the linear susceptibility of the material, which is a second rank tensor that relates the three incident field components to the polarization vector.

If we increase the strength of the electric field to an energy comparable to the energy that binds valence electrons to the nucleus of an atom, the polarizability of a material will have a more complex behaviour which can be described as a Taylor
Chapter 2. Background

series of multiple order nonlinear components

\[ P(\omega_1 + \omega_2, \omega_3) = P^{(1)}(\omega_1) + P^{(2)}(\omega_1 + \omega_2) + P^{(3)}(\omega_1 + \omega_2 + \omega_3) + \ldots \]

\(= \epsilon_0 \chi^{(1)}_{ij} E(\omega_1) + \)
\(\epsilon_0 \chi^{(2)}_{ijk} : E(\omega_1)E(\omega_2) + \)
\(\epsilon_0 \chi^{(3)}_{ijkl} : E(\omega_1)E(\omega_2)E(\omega_3) + \ldots \)  

(2.26)

where \( \chi^{(2)}_{ijk} \) and \( \chi^{(3)}_{ijkl} \) are nonlinear susceptibility tensors of rank \((n + 1)\), with \(3^{n+1}\) tensor components and units \(V^{n-1} m^{-(n-1)}\). They are physically related to the symmetry properties of the optical medium. These nonlinear polarization terms are associated with new source components of the electromagnetic field in material which is responsible for effects such as optical rectification \(\chi^{(2)}(0 : \omega, -\omega)\), sum and difference frequency generation \(\chi^{(2)}(-\omega_1 \pm \omega_2 : \omega_1, \omega_2)\), second-harmonic generation (SHG) \(\chi^{(2)}(-2\omega : \omega, \omega)\), Pockels effect \(\chi^{(2)}(-\omega : \omega, 0)\), and third-harmonic generation (THG) \(\chi^{(3)}(-3\omega : \omega, \omega, \omega)\). For higher-order effects, the nonlinear polarizations become smaller in intensity, and therefore, very strong fields are required to exploit these effects. The measurement of nonlinear optical effects was first reported in 1961 through an SHG experiment using a quartz crystal as the nonlinear medium [106]. This was made possible one year after the development of the first laser, since strong fields are required to excite nonlinear processes.

2.2.1 Relevant properties for nonlinear optical materials

The ionic structure of materials is what defines their optical properties and ability to polarize in response to an electric field. Nonlinear optical effects are strongest in materials where electrons are tightly bound to the ionic lattice. Highly ordered crystals are typically found therefore to have large nonlinear optical coefficients. These are typically quantified by the linear and nonlinear susceptibilities \(\chi^{(n)}\) in Eq. 2.26, and they are constrained by the symmetry properties of the medium [107]. Figure 2.12 shows the highest values for the second- and third-order nonlinear susceptibilities \((\chi^{(2)}\) and \(\chi^{(3)}\)) of various IR materials. Due to the high-field intensities needed to excite nonlinear processes, materials with low linear absorption and high melting points are typically adopted to avoid optical damage and alteration.
Chapter 2. Background

of their optical properties. The nonlinear optical processes scale with the number of atoms involved, and therefore, the optical transparency of a material at the fundamental and nonlinear wavelengths are an important factor to consider to build-up a stronger nonlinear signal.

![Nonlinear optical coefficients for IR materials](image)

Figure 2.12: Nonlinear optical coefficients for IR materials with (a) second- and (b) third-order nonlinearities. Values obtained from [108, 109, 110, 107, 111].

**Inversion symmetry**

Even-order terms of the polarizability in Eq. 2.26 are known to only occur in non-centrosymmetric materials, or at the surface of materials where the centrosymmetry is broken. If we consider a centrosymmetric material, the generated second-order polarization should follow the same sign as the incident field, due to the inversion symmetry of the medium, leading to \( \epsilon_0 \chi^{(2)} E^2 = \epsilon_0 \chi^{(2)} (-E)^2 \leftrightarrow P^{(2)} = -P^{(2)} \) which can only occur when \( \chi^{(2)} = 0 \), and therefore, second-order nonlinear processes are not allowed in centrosymmetric media. This is not the case however for odd-order nonlinear processes, which can occur in most media independent of the symmetry, but with various degrees of intensity depending on the penetration of the electric field inside the material.
Chapter 2. Background

Phase-matching

For efficient sum-frequency generation, the atomic dipoles generated throughout a nonlinear material need to be properly phased so that the generated field can propagate coherently in the direction of the incident field. When the individual dipoles are phased, the total radiated power scales with the square of the atoms that participate in the process.

Considering a nonlinear crystal with an effective path length \( L \) and a nonlinear susceptibility \( \chi^{(2)} \), it can be determined [107] that the intensity of the generated field from a sum-frequency process \( \omega_3 = \omega_1 + \omega_2 \) – also applicable to second-harmonic generation if we take \( \omega_1 = \omega_2 \) – is given by

\[
I_3 = \frac{\chi^{(2)}_2 \omega_3^2 I_1 I_2}{2n(\omega_1)n(\omega_2)n(\omega_3)e_0c^3} L^2 \text{sinc}^2 \left( \frac{\Delta k L}{2} \right)
\]  

(2.27)

where \( I_1 \) and \( I_2 \) are the intensities of the incident fields with frequency \( \omega_1 \) and \( \omega_2 \), \( n \) the refractive index of the crystal as a function of the incident and generated field frequencies, and \( \Delta k = k_1 + k_2 - k_3 \) the wavevector mismatch. Expanding the definition of the wavevector mismatch

\[
\Delta k = \frac{n(\omega_1)\omega_1}{c} + \frac{n(\omega_2)\omega_2}{c} - \frac{n(\omega_3)\omega_3}{c}
\]  

(2.28)

we can see that this is both dependent on the frequency of light and the dispersive properties of the nonlinear material. For the particular case of SHG \((\omega = \omega_1 = \omega_2 = \omega_3/2)\), we can obtain the phase difference \( \Delta \Phi \) between the waves at a distance \( z \)

\[
\Delta \Phi = \frac{2\omega}{c} |n(\omega) - n(2\omega)| z
\]

and then derive an expression for the coherence length \( l_c \) of the material

\[
l_c = \frac{c\pi}{\omega} \frac{1}{|n(2\omega) - n(\omega)|}.
\]

(2.29)

The coherence length corresponds to the propagation distance over which the electric-field adds coherently to build a nonlinear signal. Destructive interference of the waves occurs above the coherence length until the nonlinear signal reaches zero intensity at \( 2l_c \), from which point the same process repeats again. Typical values for the coherence length are in the order of tens of micrometers, as shown in
Figure 2.13 for gallium arsenide (GaAs) and ZnO respectively.

Figure 2.13: Coherence length for (a) GaAs and (b) ZnO materials using different incident pump wavelengths. The refractive index values used in Eq. 2.29 were obtained from [112] and [113]. The different curves in (b) correspond to the ordinary and extraordinary refractive indexes of ZnO.

Returning to Eq. 2.28, momentum conservation requires that $k_3 = k_1 + k_2$ and so $\Delta k = 0$ needs to be satisfied in order to achieve perfect phase-matching. Re-writing Eq. 2.28 so $\Delta k = 0$ and expressing it through the difference $n(\omega_3) - n(\omega_2)$ will lead us to

$$n(\omega_3) - n(\omega_2) = \frac{\omega_1}{\omega_3} [n(\omega_1) - n(\omega_2)]$$

For materials with normal dispersion, where the refractive index increases with frequency (Fig. 2.14), the equality cannot be satisfied as $\omega_1 \leq \omega_2 < \omega_3$ and so in reality we have $n(\omega_3) - n(\omega_2) > 0$ and $n(\omega_1) - n(\omega_2) \leq 0$ where $(\omega_1, \omega_2, \omega_3) \in \mathbb{N}^3$.

In theory, materials that show anomalous dispersion could potentially be used since the index of refraction drops with increased frequency (Fig. 2.14). In practice, this is not the case as the drop in the refractive index is usually associated with an absorption peak that causes it to drop and raise sharply in a narrow band of frequencies.

Phase-matching is therefore typically achieved using a class of materials with
optical anisotropy, typically crystalline solids. In contrast to single atoms, which offer the highest symmetry due to their spherical invariance, the optical properties of materials are strongly related to their translational symmetry defined by the arrangement of atoms in the materials unit cell. In anisotropic materials, we get different vibrational frequencies depending on the polarization of the optical field, which leads to a change in the refractive index - this phenomenon is called birefringence. Birefringent materials are characterised by several properties: (1) they exhibit an optical axis known as the $\hat{c}$ axis; (2) light with vector $\mathbf{k}$ propagating perpendicularly or parallel to the $\hat{c}$ axis will experience different refractive indices, i.e. ordinary refractive index ($n_o$) if $\mathbf{k} \perp \hat{c}$ or an extraordinary refractive ($n_e$) if $\mathbf{k} \parallel \hat{c}$. Control of the refractive index in anisotropic crystals can be achieved through changing the temperature of the crystal, or adjusting the optical axis of the crystal in respect to the direction of propagation of the incident light. In both methods, birefringence is exploited by making the generated optical field with highest frequency ($\omega_3 = \omega_1 + \omega_2$) to be polarized in the direction of the crystalline axes that gives the lowest refractive index. A variety of birefringent materials
typically used in second-order nonlinear applications are highlighted in Fig. 2.12.(a).

There is a class of isotropic materials with much larger $\chi^{(2)}$ values, as seen in Fig. 2.12.(a), which have only limited applications in second-order nonlinear processes due to their lack of birefringence. These are typically of the cubic crystal structure, thus have no preference on the direction of the field perturbations and cannot be phase-matched. This issue can be overcome with quasi-phase matching by using a periodically poled material, where the orientation of the crystalline axes is inverted at a distance equal to the coherence length which allows constructive interference of the nonlinear signals in large lengths of isotropic or low birefringent materials [114, 115]. Alternatively, one can also use built-in artificial birefringence to break the isotropy of a nonlinear material [116, 117], or a modal phase-matching method for waveguide nonlinear generation [118, 119]. In Chap. 5, we will discuss an alternative method for efficient nonlinear generation with isotropic or low birefringent materials, using electric field enhancements from plasmonic nanostructures.

### 2.2.2 SHG effects in centrosymmetric media

Previously, we have seen that in the electric dipole approximation all nonlinear optical effects of even orders cannot occur in centrosymmetric materials. Alternatively, even-ordered nonlinearities can be observed at the interface between materials where the inversion symmetry is broken [120]. The surface second harmonic polarization $P_s^{(2\omega, r)}$ of isotropic and centrosymmetric media at position $r$ can be described as [121]

$$P_s^{(2\omega, r)} = \epsilon_0 \chi_s^{(2)} : E(\omega, r)E(\omega, r)\delta(r - r_S)$$

(2.30)

where $\chi_s^{(2)}$ is the surface second order nonlinear susceptibility, $\delta$ the Dirac delta function, and $r_s$ the surface vector. The last two terms are used to define a angstrom thin layer representing the surface of the material. In isotropic and centrosymmetric materials, as in the case of metals, the surface second order nonlinear susceptibility consists of only three nonvanishing elements ($\chi_{\parallel\parallel}^{(2)} = \chi_{\parallel\perp}^{(2)}$ and $\chi_{\perp\perp\perp}^{(2)}$) which are
associated with surface currents oscillating perpendicular and parallel to the surface [122]. Surface SHG is then the results of two polarization components described as

\[ P_{\perp}^{s}(2\omega, r) = \epsilon_0 \chi^{(2)}_{\perp\perp} E_{\perp}^2(\omega, r), \]  
\[ P_{\parallel}^{s}(2\omega, r) = 2\epsilon_0 \chi^{(2)}_{\parallel\parallel} E_{\parallel}(\omega, r) E_{\perp}(\omega, r). \]  

Adding to the surface effects, isotropic and centrosymmetric media can also support a second contribution to the SHG derived from magnetic dipole and electric quadrupole effects. The formula for SHG polarization which accounts for bulk effects is given by [123]

\[ P_{b}^{i}(2\omega, r) = \gamma \nabla_i(\mathbf{E} \cdot \mathbf{E}) + \delta(\mathbf{E} \cdot \nabla)E_i + \beta(\nabla \cdot \mathbf{E})E_i + \zeta E_i \nabla_i E_i \]  

where \( \mathbf{E} = \mathbf{E}(\omega, r) \) is the fundamental electric-field inside the material at position \( r \), and \( \gamma, \delta, \beta \) and \( \zeta \) are material parameters, the first three are associated with an isotropic response (independent of sample orientation), and \( \zeta \) with a bulk anisotropic response. For isotropic and centrosymmetric media such as metals, the last term is neglected (\( \zeta = 0 \)). For a homogeneous field distribution in a continuous metallic film, the first term can also be neglected since according to Maxwell’s equations the term \( \nabla \cdot \mathbf{E} = \nabla \cdot \mathbf{D}/\varepsilon \) vanishes.

Combining the surface and bulk contributions, the total SHG response of a isotropic and centrosymmetric media is described by [121]

\[ \mathbf{P}(2\omega, r) = \mathbf{P}^{s}(2\omega, r) + \mathbf{P}^{b}(2\omega, r). \]  

In Chap. 5, we will discuss in more detail SHG effects from chiral plasmonic nanostructures. In this metamaterials, the local fundamental fields are highly inhomogeneous, and the SHG signal can be enhanced through localized plasmonic effects and superchiral fields.

### 2.3 Applications of nonlinear plasmonics

Since the observation of second-harmonic generation (SHG) in 1961, the field of nonlinear optics has evolved from fundamental studies of light-matter interactions
to the development of new technologies. One of the key features of this topic is the possibility of developing techniques which enable the control of light with light, by means of frequency and optical field intensity. Nonlinear frequency conversion and in particular the optical parametric oscillator (OPO) [10], are possibly the most widely adopted nonlinear technology, allowing extension of the wavelength range of a laser with conversion efficiencies in the range of tens of percent. Another area of interest is in all-optical switching, where strong field intensities are used to induce a change in refractive index of a nonlinear material, and deflect or delay light signals. The switching speeds are typically in the order of 100ps, making these systems extremely useful in optical high-speed communications [124, 125]. Sensing is also another area where nonlinear effects are used extensively as a mean for a nondestructive, in situ and real-time monitoring of various biological, chemical and physical processes [126, 127].

A key limitation of modern nonlinear systems is the large field intensities required to excite inherently weak optical signals, and as such, large and relatively complex optical systems are required to attain measurable results. This in turn is a major bottleneck for the development of integrated photonic systems, where size and power are costly requirements. Fortunately, some solutions can be found in plasmonics such as the field enhancement and geometric confinement obtained from surface plasmon excitations, improving of efficiency and size of existing nonlinear techniques. In the area of frequency conversion, plasmonic enhancements have been used for bulk SHG in isotropic nonlinear media [16], high-harmonic generation from an argon gas jet [128], SHG from isotropic and anisotropic plasmonic crystals [129, 130], and SHG from magnetic resonant metasurfaces [131].

Optical effects from plasmonic nanostructures are also sensitive to geometrical transformations of the metallic surface (e.g. rotation and mirror symmetry). This property has been exploited in areas requiring symmetry-selective harmonic generation [132, 133], nonlinear circular dichroism and optical activity [134, 71] and nonlinear phase control [135]. Furthermore, as we have seen in Sec. 2.1.3, the excitation of SPPs is largely influenced by the dielectric interface, which in conjunction
with the concentrated electric-field at the surface of metallic nanostructures and the higher contrast from nonlinear effects, serves as a sensitive probe for various processes such as the assessment of the chiral purity of molecules [70], refractive index sensing [136, 137, 138], nano ruler [139, 140] and near-field imaging [141].
Chapter 3

Experimental and numerical methods

The first section of this chapter is dedicated to the fabrication techniques and processes of planar nonlinear metasurfaces. We describe individually the main techniques used for nanostructuring a range of metallic films and substrates, including discussion of ideal process conditions and limitations in the optical response. We will then conclude the first section with a review of the different process flows for producing a range of metasurfaces. In the second section, we introduce the techniques that were used to analyse the linear and nonlinear optical response from fabricated devices. In particular, the experimental requirements for measurement of highly sensitive and weak signals in the near- to mid-IR range. Finally, we describe the main numerical methods used in this work to model the optical response of different devices. In particular we describe the rigorous coupled wave analysis (RCWA) method, used in the calculation of the spectral response and near-field distribution at the fundamental frequency, and a multipole expansion of the linear and nonlinear polarizations, to determine the total and individual contributions of the transmitted signals.
3.1 Fabrication techniques

In this section, we will start by discussing individually the main techniques used in fabrication of nonlinear plasmonic samples. At the end (Sec. 3.1.7), we will integrate all this information and summarize the main process flows for producing complementary plasmonic structures of arbitrary shapes, based on the presented techniques.

3.1.1 Substrate preparation

Throughout this work we have used a range of substrate materials based on devices requirements, the most common being:

- GaAs – (100) oriented, undoped (semi-insulating), double side polished, 3 inch wafers from Wafer Technology Ltd;
- Al$_2$O$_3$ – (0001) c-plane oriented, 2 inch wafer, single side polished substrates from SurfaceNet GmbH;
- Silicon (Si) – doped, thermally oxide ($\approx 100$nm), single side polished, 3 inch wafer.

All wafer substrates were diced using a Disco DAD3230 dicing saw to make $7.5 \times 7.5$ mm samples. All materials are optically transparent in the wavelength of interest, with the exception of Si which was used in reflection measurements, and the oxide layer introduced to blue-shift the plasmonic resonance to a detectable range ($\varepsilon_r$(SiO$_2$) $< \varepsilon_r$(Si)). Before processing, the substrates were thoroughly cleaned to remove dust particles and organic contaminants sitting at the surface. This was achieved by using a degreasing process based on a sequence of treatments using organic solvents [142]: 1) 10 min in a ultrasonic bath of boiling acetone, 2) 10 min soaking in methanol, and 3) 10 min soaking in isopropyl alcohol (IPA). The samples where then rinsed in dionized water and dried with a N$_2$ gun. Any residual organic materials left from the solvent clean were removed using an O$_2$ plasma ashing step for 10 min at 300W and 0.4 mbar. Because of the low ion energies, with respect to the atomic-displacement threshold of the substrate materials, the ashing
process has no damaging effect on the material structure or plasmonic response.

3.1.2 Electron-beam lithography

Electron-beam lithography (EBL) has been used extensively in this work to pattern arbitrary planar nanostructures into an electron-sensitive film, which was then further processed to transfer the features into a range of materials (Sec. 3.1.7). In contrast to photolithography systems, which already deliver resolution down to 30 nm thanks to advancements in deep-UV optics, large numerical aperture lenses and computational methods for mask design, electron-based systems can only achieve a slight improvement in resolution. In principle, given the shorter wavelength nature of electrons in comparison to photons, one would expect significant resolution gains from electron-based systems. However, due to aberration effects in the EBL lenses and limited sensitivity of the e-beam resist, the smallest attainable feature is typically on the order of a few nanometers. The strength of EBL systems are therefore in the simplicity and flexibility to fabricate one of a kind samples, without incurring large costs to produce a single optical mask, making it an indispensable tool in research and development laboratories.

Electron column

The working principles of EBL are closely related to a scanning electron microscope (SEM), in that an electron column is used to generate, focus and scan a beam of electrons onto a sample in a highly controlled manner. In Fig. 3.1 is shown a cross-section schematic of a Carl Zeiss Gemini electron column, the main component of the RAITH150 Two EBL system used in this work. For this particular column, a thermal field emission source is used to generate a high current/brightness beam of electrons. To achieve this, a tungsten tip and suppressor electrode are negative biased at several kV relative to a nearby extractor electrode, leading to a high potential gradient at the surface of the tip and tunnelling of electrons from the gun. The tunnelling distance is minimized by applying some thermal energy to the
system, making the electron extraction process more efficient and increasing the lifetime of the tungsten tip. The beam diameter of field emission sources are in the 5 to 10 nm range. In comparison to alternative electron sources (e.g. thermionic or cold field emission), thermal field emission offers the best compromise in terms of beam stability, brightness, beam diameter, energy spread, lifetime and vacuum requirements [143].

After the electron source, the beam goes through a set of lenses, deflection systems and apertures to demagnify the beam by one or two orders of magnitude. The ultimate resolution is defined by the diffraction limit of the column, dependent on the electron wavelength and beam aperture angle. However, spherical, astigmatic and chromatic aberrations largely compromise the final spot diameter, due to imperfections in the lenses and energy spread from the source. The deflection systems in the column are used to control the X-Y position of the beam in the sample, which is used for scanning purposes during imaging, and for patterning
arbitrary structures on a electron sensitive resist.

**Electron-solid interaction and detection**

A range of electron-solid interactions occur when the electron-beam reaches the sample. Of particular interest for EBL and SEM imaging applications are forward and backward scattering events. The former are associated with high energy losses (inelastic) from bound electron interactions within the sample, leading to the generation of (lower energy) secondary electrons with a small propagation length. Backward scattering effects tend to occur further inside the sample with a larger deviation angle, and are the result of atomic nucleus interactions, leading to reflection of the incident electrons with a small loss of energy. Information regarding the surface topography or element composition can therefore be obtained by looking at different scattering events.

Detection of secondary and backscattered electrons requires different strategies due to the difference in energy. Typically this achieved using a Everhart-Thornley electron detector, consisting of a scintillator, which converts the detected electrons to a light or electronic signal, and a Faraday cage, that draws low-energy secondary electrons towards the scintillator. Backscattered electrons can be detected without the Faraday cage as long as their direction is in line of sight with the scintillator. However, due to the high energy content, backscattered electrons can undergo multiple reflections inside the specimen chamber, leading to similar electron-matter interactions as in the sample (i.e. conversion to other backscattered or secondary electrons), and introduction of noise in the detection side. A way around this is to integrate a (in-lens) detector inside the electron column, where the lens works as a filter preventing the backscattered electrons from having a line of sight to the detector, and producing a pure secondary electron image. The SEM imaging systems used in this work, Leo XB1540 cross-beam and RAITH150 Two, are both equipped with an in-lens and specimen chamber Everhart-Thornley detectors.
Electron-beam resists

An EBL process requires deposition of an electron sensitive resist on top of the sample prior to exposure. The electron beam effect on the resist occurs through fragmentation or cross-linking of molecular chains, making the resist more (positive-type) or less (negative-type) soluble in a developing solution. Typically, forward scattering is the most dominant effect in the exposure of electron-beam resist, however, backscattering can also play a role, in particular for larger acceleration voltages. Due to the large propagation length of backscattered electrons, proximity effects can be problematic in the exposure of closely spaced features.

In this work, we have used poly(methyl methacrylate) (PMMA) from MicroChem Corp has a positive resist. Due to its polymer based properties, PMMA is highly soluble in solvent, allowing it to be easily integrated in a range of semiconductor processes. Other advantages include the low cost, chemical stability and process reliability, making PMMA the standard positive electron beam resist. In terms of resolution, PMMA is however not the best performing resist on the market, given the large polymer size in the order of 5 to 10 nm. In our EBL system we have achieved a minimum feature size of 25 nm on a 100 nm resist, as highlighted by processes (a) and (b) in Tab. 3.1 and Fig. 3.2. Removal of the fragmented polymer chains (a.k.a. development) can be achieved using a dilution of methyl isobutyl ketone (MIBK) with IPA solvents (typically a 1:3 ratio). The full removal of the resist is done using acetone or O\textsubscript{2} plasma ashing. An interesting property of PMMA is that by applying a sufficiently large electron exposure dose it is possible to cross link the polymer chains and change the tonality of the resist – illustrated in Tab. 3.1.(c) and Fig. 3.2.(c). This is however subject to large proximity effects which limit the process resolution.

A better resolution can be achieved with an inorganic negative resist known as hydrogen silsesquioxane (HSQ), manufactured by Dow Corning\textsuperscript{®} with the trade name XR1541. In our EBL system, we have achieved feature sizes down to 10 nm using a very thin film of HSQ, diluted in MIBK from a 2% bottle to 0.5%, as
shown in Tab. 3.1.(e) and Fig. 3.2.(e). After exposure and development, HSQ has a chemical composition similar to SiO$_2$ making it difficult to dissolve using solvents. Instead, removal is achieved through an etching process, potentially limiting the applicability of the resist. In Sec. 3.1.7, we will discuss an alternative bi-layer resist method, combining PMMA and HSQ, where some of this limitation can be overcome.

<table>
<thead>
<tr>
<th>#</th>
<th>Resist / Thickness</th>
<th>Substrate</th>
<th>EBL parameters</th>
<th>Developer / Time</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>950 PMMA A2 100 nm</td>
<td>doped Si</td>
<td>30 kV 20 $\mu$m 0.396 pA</td>
<td>MIBK:IPA (1:3) / 60 s</td>
<td>dot size 25 nm period 50 nm</td>
</tr>
<tr>
<td>(b)</td>
<td>950 PMMA A2 100 nm</td>
<td>doped Si</td>
<td>30 kV 20 $\mu$m 792 $\mu$As/cm</td>
<td>MIBK:IPA (1:6) / 60 s</td>
<td>line width 25 nm spacing 50 nm</td>
</tr>
<tr>
<td>(c)</td>
<td>950 PMMA A2 100 nm</td>
<td>undoped GaAs</td>
<td>10 kV 30 $\mu$m 1100 $\mu$As/cm$^2$</td>
<td>Acetone / 60 s</td>
<td>feature 80 nm</td>
</tr>
<tr>
<td>(d)</td>
<td>XR1541 2% 100 nm</td>
<td>doped Si</td>
<td>10 kV 20 $\mu$m 220 $\mu$As/cm</td>
<td>NaCl:NaOH / 60 s</td>
<td>line width 11 nm spacing 24 nm</td>
</tr>
<tr>
<td>(e)</td>
<td>XR1541 0.5% 100 nm</td>
<td>doped Si</td>
<td>10 kV 20 $\mu$m 51 $\mu$As/cm</td>
<td>MF-26A / 60 s</td>
<td>line width 10 nm spacing 10 nm</td>
</tr>
</tbody>
</table>

Table 3.1: Selection of EBL experiments using different process parameters.

Figure 3.2: SEM micrographs of various resist structures based on the EBL parameters presented in Tab. 3.1.

In Fig. 3.3, the effect of electron dosage on the plasmonic response of a slit aperture is presented, showing no significant deviation from the simulation result as illustrated by the dashed line. In this exercise, argon milling (Sec. 3.1.7) was used to transfer the patterns from the exposed PMMA to the metallic film. For the lowest dosage, a slight attenuation of the transmitted signal is observed, due to
insufficient exposure of the PMMA leading to a shallow mill of the metallic film. Slight changes in the aperture size occur for increased dosage values, which explain the slight variations in the different curves.

Figure 3.3: FTIR transmission measurement of an array of plasmonic slit apertures fabricated using the EBL at 10 kV acceleration voltage, 30 $\mu$m aperture, and varied electron dosage (73 to 165 $\mu$C/cm$^2$). Structures were milled using argon milling for 15 min at $V_{\text{acc}}=390$ V, $I_b=10$ mA and $V_b=200$ V. Substrate material is CaF$_2$, and thickness of metallic films 30 nm gold and 5 nm titanium. FTIR measurement using linear polarized light, with electric-field vector orthogonal to the long dimension of the slits. The images in the dose labels correspond to a SEM micrograph of a single slit, fabricated with the specified dosage. The dashed line shows the simulated result obtained using a commercial implementation of RCWA (Sec. 3.3.1). The model was implemented using $N=8$ harmonics and a fixed refractive index value of 1.39 for the CaF$_2$ substrate.

**Focusing electron column**

For reproducible results, it is important to achieve a consistent alignment, focus and spot size with the electron column. Typically, an EBL user relies on small particle
defects in the sample to achieve a good alignment, e.g. a small scratch in a corner of the sample using a scriber. However, sometimes it is difficult to find particles of small dimension to achieve a good focus. An alternative method is to use a colloidal solution of gold nanoparticles, with sizes around 25 nm or below. These can be easily purchased from various suppliers and come dispersed in deionized water. A small drop of this solution can be deposited on a corner of the sample using a pipette after spin-coating the e-beam resist, and then dried-off during the resist baking.

3.1.3 Focused ion-beam milling

Focused ion-beam milling (FIB) is an alternative technique to EBL, which can be used to nanostructure directly a range of materials, without the need of extra processing steps. The working principles of an FIB column is closely related to an SEM, with the main difference being the use of high-energy ionized atoms rather than electrons. These have a greater mass and momentum than electrons, making them extremely useful in milling applications [144]. Ions also have a wavelength smaller than electrons of the same energy, and in principle, can offer greater resolution gains if only diffraction limited aberrations are considered. Similarly to an SEM, ions undergo elastic and inelastic scattering effects. Elastic/high-energy effects are mostly responsible for the sputtering of atoms from the surface of the sample, while inelastic/low-energy collisions lead to a transfer of energy, and the emission of secondary electrons which can be captured for imaging purposes. Image resolution of an FIB can supersede the SEM, and is typically limited by the differential rate of sputtering and secondary electron detection [145]. One can use lower mass ions to minimize the sputter yield (i.e. number of sputtered atoms versus incident ions) and improve resolution. An example of this, is the helium FIB which can achieve sub-nanometer resolution.

In this work, FIB was used in conjunction with SEM to mill and image 3D cross-sections on various samples, which helped to identify problems in fabrication. The main FIB instrument was a Leo (now part of Carl Zeiss) XB1540 cross-beam,
Figure 3.4: (a) FTIR transmission measurement of an array of plasmonic slit apertures, (b) transmission versus dose plot for three wavelengths, corresponding to features in the transmission data. Samples were fabricated using a gallium FIB at 30 kV acceleration voltage, 50 pA beam current, and varied ion dose (0.01 to 0.1 nC/cm$^2$). Substrate material is CaF$_2$, and thickness of metallic films 30 nm gold and 5 nm titanium. FTIR measurement using linearly polarized light, with electric-field vector orthogonal to the long dimension of the slits.

which is a gallium based system with a Orsay Physics FIB column and a Gemini SEM column. Gallium is a highly common element in these instruments due to the low melting point and low volatility during operation, allowing an efficient use of material while prolonging the lifespan of the source tip. Also the gallium mass offers a good compromise between sputtering yield and sample integrity, even for samples with heavier elements. Imaging with gallium FIB is possible but very limiting in terms of resolution due to the large sputter yield even for low energy ions. This is why these systems are integrated with an SEM column allowing simultaneous imaging during milling. Another aspect of gallium FIB is the large implantation of ions during milling which might influence the electrical and optical properties of the sample. In Fig. 3.4 is shown the effect of gallium implantation on the optical response of slit apertures in a gold film deposited on a calcium fluoride (CaF$_2$) substrate. The structures were milled using gallium FIB at different ion dosages, showing that the strength of the plasmonic resonance is influenced by the depth of the mill. There is poor transmission of light for low ion-beam dosage and shallow mills. However it reaches a stable level once the metallic film is removed completely.
from the aperture, occurring for dose above 0.06 nC/cm². The peak transmission is slightly lower than the result obtained from the EBL experiment, likely due to the gallium implantation. Also, a feature appears at 6.29 µm, which could be explained by an anomalous dispersion of the gallium at that wavelength.

3.1.4 Argon ion milling

Ion milling was used in conjunction with EBL to transfer the patterns from the e-beam resist to a metallic film or substrate. It is a physical etching process where a collimated beam of ions is accelerated towards a sample, and through kinetic energy it slowly sputters material from the surface. Ions are formed using elements with filled energy levels (inert gases), therefore avoiding any chemical reaction with atoms at the surface of the target sample. Argon gas is commonly used in this systems since it is both cheap and easy to obtain (forming 0.93% of the global atmosphere), and it is a sufficiently heavy element to be used in milling applications, without significant damage to the sample.

![Figure 3.5: Schematic diagram of a simplified version of an argon ion-miller. Image adapted from [146].](image)

Figure 3.5 shows a simplified schematic representation of a ion-milling system. In general, the main event occurs in a discharge chamber, where there is a constant flow of argon gas which becomes ionized through thermionic emission from the
cathode filament. Through the application of a discharge voltage between the cathode and the anode, and a magnetic field around the discharge chamber, the electrons emitted from the cathode undergo many collisions with the argon atoms leading to a more efficient ionization process. Ionized argon atoms reaching the screen grid are then accelerated towards the sample through a potential difference between the screen and acceleration grids. The main function of the screen is to minimize the wear of the acceleration grid. An extra filament near the accelerator grid is then used for thermionic emission of electrons, served to neutralize the ionized beam and reduce charging effects at the sample.

<table>
<thead>
<tr>
<th>Material</th>
<th>Acceleration voltage (V)</th>
<th>Beam current (mA)</th>
<th>Beam voltage (V)</th>
<th>Argon flow (sccm)</th>
<th>Milling rate (nm/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold</td>
<td>390</td>
<td>10</td>
<td>200</td>
<td>5</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>390</td>
<td>9</td>
<td>200</td>
<td>5</td>
<td>3.3</td>
</tr>
<tr>
<td>GaAs</td>
<td>390</td>
<td>10</td>
<td>200</td>
<td>5</td>
<td>5.5</td>
</tr>
<tr>
<td>HSQ</td>
<td>390</td>
<td>10</td>
<td>200</td>
<td>5</td>
<td>1.3</td>
</tr>
<tr>
<td>PMMA</td>
<td>390</td>
<td>10</td>
<td>200</td>
<td>5</td>
<td>3.4</td>
</tr>
<tr>
<td>Nickel</td>
<td>390</td>
<td>9</td>
<td>200</td>
<td>5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 3.2: Argon milling conditions and rates for different materials using a SVS 6000 sputtering system. Base pressure below $10^{-6}$mbar and sample rotation 58rpm.

In this work we have used a V6000 sputter system from Scientific Vacuum Systems Ltd, with an integrated ion beam milling facility. In Tab. 3.2, the milling rates and process parameters for different materials using this system are presented. For consistency of results, the system was pumped to a base pressure below $10^{-6}$mbar to minimize the number of particles in the chamber, and the sample rotated at a speed of 58rpm to prevent shadowing effects during the milling. Also, the cathode film was warmed up for 10min to ensure stability of the ion beam during the process.

The filaments and grids on the SVS V6000 milling system are made of tungsten and molybdenum respectively. In principle, there is a small chance of contaminating
Chapter 3. Experimental and numerical methods

Figure 3.6: FTIR transmission measurement with linear polarized light, of an array of plasmonic slit apertures fabricated using the EBL at 10 kV with a 30 µm aperture and a dose of 128 µC/cm², and varied milling times at $V_{\text{acc}}=390$ V, $I_b=10$ mA and $V_b=200$ V. Substrate material is CaF₂, and thickness of the films 30 nm gold and 5 nm titanium. Incident electric-field polarization with vector orthogonal to the long dimension of the slits.

3.1.5 Plasma ashing

A plasma asher was extensively used in this work to remove traces of organic contamination, create surface terminated oxygen interfaces, and for selective patterning of PMMA films. Plasma ashing works by generating a plasma of highly reactive atoms and radicals from an oxygen gas source. In contrast to other plasma based dry etching processes (e.g. reactive ion etching), plasma ashing does not force directional movement of the reactive atoms, and therefore, only low
energy ions reach the sample. As a result, very low lattice and surface damage is expected for processes with energies below the atomic displacement energy of the sample. A chemical reaction occurs when the reactive species interacts with the organic polymer in the sample, forming gaseous organic molecules (ashes) which are removed from the system by the vacuum pump.

In this work we have used a Diener plasma asher with a constant base pressure of 0.4mbar (set by the oxygen gas line). A typical cleaning routine involved a 10min process at 300W power. For controlled patterning of the PMMA using a bi-layer resist process (Sec. 3.1.7), we have determined an ashing rate of 11nm/min at 200W power. The time of the process was adjusted based on the thickness of the resist.

3.1.6 Thermal evaporation

Thermal evaporation is a common technique used to deposit a range of materials, in particular metallic films, due to its simplicity and reliability. The working principles for evaporation in presented in Fig. 3.7 and, as the name suggests, it works by heating a material until melting point, causing it to evaporate from the source and condense at a sample stage, forming an amorphous film with a surface roughness of a few angstroms. There are essentially two methods to achieve melting and evaporation of the target material: (1) resistive evaporation, which relies on using a current source connected to a resistive element with higher melting point than the target material; (2) electron beam evaporation, which directly heats the target material to melting point, using a focussed beam of high energy electrons [147]. Electron beam evaporation tends to be used for depositing materials with a larger melting temperature, and it gives films with higher purity and lower surface roughness than thermal evaporation.

As illustrated in Fig. 3.7, evaporation is achieved in line of sight from a point source (heating filament or crucible) located at a distance d from the sample. This typically results in poor step edge coverage and non-uniformity of the film thickness across large areas. The kinetic energy associated with thermal evaporation is usually
very low, and therefore no physical damage to the sample surface is expected. However, the effect of high thermal radiation on the sample might have to be considered, in particular when using polymer based resists. Faster evaporation rates and/or cooled sample stages can be used to minimize the effect of thermal radiation on the sample. Alternatively, one should consider sputtering for lower temperature material deposition, with the added advantage of improved step coverage and uniformity, but at a cost of system complexity, equipment and running costs, poorer film quality (prone to defect creation and argon implantation) and lower deposition rates.

In thermal evaporation, high vacuum is required to increase the mean free path of the source material and minimize contamination of the sample with other particles in the vacuum chamber. Ideally, one wants the time for a residual vacuum material to reach the sample to be much longer than that of the evaporated material. This is achieved by setting the distance $d$ between source and sample stage to be smaller than the mean free path, typically 0.2–0.5m to avoid radiative heating, and by reducing the base pressure of the system, ideally below $10^{-6}$mbar. Contamination is also minimized by limiting the number of materials that can be evaporated in a single system. In particular one needs to avoid mixing magnetic and non-magnetic materials.
Chapter 3. Experimental and numerical methods

<table>
<thead>
<tr>
<th>Material</th>
<th>Melting point [°C]</th>
<th>Melting point [°C] @ 10^-7 mbar</th>
<th>Density [g cm^-3]</th>
<th>Z-ratio [Auto 206]</th>
<th>Z-ratio [Auto 500]</th>
</tr>
</thead>
<tbody>
<tr>
<td>chromium</td>
<td>1857</td>
<td>837</td>
<td>7.2</td>
<td>28.94</td>
<td>–</td>
</tr>
<tr>
<td>gold</td>
<td>1064</td>
<td>807</td>
<td>19.32</td>
<td>23.17</td>
<td>0.381</td>
</tr>
<tr>
<td>titanium</td>
<td>1660</td>
<td>1067</td>
<td>4.5</td>
<td>–</td>
<td>0.628</td>
</tr>
</tbody>
</table>

Table 3.3: Relevant material parameters for thermal evaporation. Extracted from the material deposition chart supplied by Kurt J. Lesker.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Material</th>
<th>Holder</th>
<th>Parameters</th>
<th>Rate [nm/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auto 306</td>
<td>Cr</td>
<td>plated tungsten rod</td>
<td>10V 4.2A</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Au</td>
<td>molybdenum or tungsten boat</td>
<td>10V 4.6A</td>
<td>0.3</td>
</tr>
<tr>
<td>Auto 500</td>
<td>Ti</td>
<td>none</td>
<td>5kV 40mA</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Au</td>
<td>tungsten crucible</td>
<td>5kV 50mA</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Table 3.4: Process parameters for thermal evaporation of chromium, titatium and gold using a Edwards Auto 306 (resistive) and Edwards Auto 500 (electron beam) evaporators.

All devices in this work operate in the near- and mid-IR range, where gold is the best material in terms of low optical losses, high reflectivity and good plasmonic quality factors [30]. Gold is however known for poor adhesion properties, requiring deposition of an intermediate layer of an oxidative metal, such as chromium or titanium, to improve adhesion to the substrate. Evaporation of the adhesion layer is done before deposition of the gold without breaking the vacuum, and a thickness of 5 nm is usually sufficient to obtain good results without compromising the plasmonic response of the gold film.

In this work an Edwards Auto 306 resistive thermal evaporator was used for deposition of gold and chromium films, and an Edwards Auto 500 electron evaporator for deposition of gold and titanium. Both instruments were operated
at a base pressure below $10^{-6}$ mbar, and the evaporation rate monitored using an in-built monitor system. Essentially the monitor measures the oscillation frequency of a quartz crystal during the process, this being dependent on the mass density and rate of the evaporated material. Correct configuration of the monitor requires knowledge of the material density and the acoustic Z-ratio, parameter that accounts for frequency change with evaporated thickness. These values are presented in Tab. 3.3. Further information on process parameters (supply voltage and filament/electron beam current), holder types and evaporation rates are presented in Tab. 3.4. The individual material and instrument tooling factors, which correct for differences between the real and estimated thickness values, were obtained from a log sheet based on the historical performance of the instrument. Other details include having the sample positioned at the centre of the sample stage, to prevent a non-uniform coating of the film, and avoiding the use of a rotating stage during the process. The latter is of relevance for lift-off processes as it minimizes step-coverage.

3.1.7 Fabrication processes for plasmonic metamaterials

These follows now a discussion of different process flows to produce a range of plasmonic nanostructures. Figure 3.8 shows the process flow for five different fabrication methods, of which we will now give an overview and discuss the trade-offs between each method. Since the process parameters vary considerably depending on the device specification, we will leave the full details of the recipe for later sections.

**Process (A,B) - bi-layer EBL resist**

Processes (A,B) are based on a negative lift-off method [148], with the added step in Process (B) for milling or etching the substrate, allowing the plasmonic apertures to be filled with the substrate material. This method has been used for fabrication of nonlinear devices using GaAs, as shown in Appendix ???. The method uses a bi-layer of PMMA and HSQ e-beam resists to achieve the negative lift-off. This is because HSQ is insoluble in solvent, and typically, can only be removed
Figure 3.8: Sample process flow for two plasmonic configurations based on a range of fabrication techniques. In processes (B,G) the plasmonic devices are etched into the substrate, while for processes (A,C-F) the nanostructures are sitting on the surface.
with a hydrofluoric based acid or a CF$_4$/H$_4$ RIE process, which might damage the substrate material. A polymer based resist, such as PMMA, is then used underneath the HSQ to achieve a solvent based lift-off. After EBL exposure and HSQ development, the patterns are transferred to the PMMA through O$_2$ plasma ashing, followed by an anisotropic etch or mill of the substrate. A good lift-off is achieved with a slight undercut of the PMMA resist during the plasma ashing step. However, for very small nanostructures this becomes difficult to control and typical minimum feature sizes are in the range of hundreds of nanometers, limiting the wavelength response of the nanostructures to mid-IR. Another disadvantage is that due to the EBL exposure, the process is limited to conductive substrates. For semi-insulating (undoped) substrates, charging effects can be minimized by coating the HSQ resist with a few nanometers gold film or PEDOT:PSS water soluble polymer, before the EBL exposure. Sputtering of gold should be avoided in this step, as it will result in cross-linking of the HSQ film, and thermal evaporation should be used instead.

**Process (C) - direct milling of nanostructures using FIB**

Process (C) is the simplest of all, as it requires only three steps, and it can be used with non-conductive substrates. It uses FIB to mill the nanostructures directly into the metallic films, and therefore is a quick and cheap technique for prototyping. A downside is the ion implantation discussed in Sec. 3.1.3 which affects the plasmonic response and losses in the metal. Also the milling times can scale significantly with the sample area and film thickness, limiting the applicability of this technique in a production environment.

**Process (D-E) - EBL and argon milling of plasmonic films**

An alternative to FIB, which avoids the ion contamination, consists of an EBL process followed by an argon milling of the metallic film. Argon is an inert and stable gas, and when used for milling causes very limited damage to the sample, with no significant alteration of the plasmonic response. Another advantage is that
the process can be easily adapted to fabricate different metamaterial configurations – i.e. aperture/transmissive or opaque/reflective structures – by using positive or negative e-beam resists (Fig. 3.8), and is also compatible with non-conductive substrates. Attention has to be taken with the milling times so as to avoid milling the substrate. Also, when using HSQ resist, one might want to consider the effect of the resist stripping step on the substrate – typically hydrofluoric acid or \( \text{CH}_4/\text{O}_2 \) RIE. In case of incompatibility, there is the possibility of implementing a bi-layer resist such as in Process (A), where the final step can be simply achieved with a solvent based process.

**Process (F-G) - EBL, argon mill of substrate and lift-off**

Processes (F-G) are essentially identical to each other, with the main difference being an added argon milling step in Process(F) to embed the nanostructures inside the substrate. Instead of milling the metallic films are deposited after the EBL process, and a solvent based lift-off is used to nanostructure the films. Chances of contamination of the films are reduced using this process and the samples give the best plasmonic response. These processes are however limited to conductive substrates, to avoid charge build-up problems using the EBL. Patterning on undoped (semi-insulating) substrates can be done by coating the PMMA resist with a few nanometers thin gold film using a basic gold sputterer. A variation of process (G) has been used in Appendix ?? to produce plasmonic metastructures embedded in a GaAs substrate.

### 3.2 Optical characterization

#### 3.2.1 Fourier transform infra-red spectroscopy

The Fourier transform infra-red spectroscopy (FTIR) technique was used extensively in this work to determine the linear spectral response of fabricated devices. A Bruker IFS66v/S FTIR was used in conjunction with a Bruker IRscope II microscope for transmission and reflection spectral measurements on small samples.
A simplified diagram of this instrument is shown in Fig. 3.9. It essentially consists of a Michelson interferometer [149] composed of four arms, containing a globar source, beam-splitter, and fixed and movable mirrors. The globar source is a thermal light source that emits unpolarized radiation of multiple wavelengths in the IR spectrum (1-20 \( \mu m \)). Light is then collimated using a collimated mirror, and directed towards a potassium bromide (KBr) beam-splitter, with transparency between 1.5 to 27 \( \mu m \). The function of the beam splitter is to split the IR radiation into two optical beams of equal intensity towards a fixed mirror and a movable mirror. The reflected beams from the mirrors are then subject to constructive and destructive interference at the beam-splitter, and the signal is re-directed towards the detector.

An actuator is used to move the movable mirror back and forward with a constant speed, and its position is recorded using a HeNe laser in conjunction with the detected signal. An interferogram is formed after completing a scan, corresponding to a full cycle of the movable mirror, which is a measurement of light intensity as function of the optical path difference between mirrors (known as retardation, \( \delta \)). Figure 3.10.(a), shows an example of a real interferogram obtained from an average of 800 scans. The two centerbursts are the result of strong constructive interference of IR light, each corresponding to an interferogram collected at different directions of the movable mirror.

**Principles of a Michelson interferometer**

The workings of a Michelson interferometer are best understood by considering the simple case of a monochromatic source with (spectroscopic) wavenumber \( \tilde{\nu} = 1/\lambda \), where \( \lambda \) is the source wavelength, and a moving mirror that can move an infinitely long distance \( (0 \le \delta < \infty) \). In this arrangement, constructive interference of the reflected waves at the beam-splitter occurs for integer multiples of the source wavelength \( \delta = n\lambda \) where \( n \in \mathbb{N}_0 \), and likewise, destructive interference occurs for retardation values of \( \delta = (n + 1/2)\lambda \). This effect can be described by a simple sinusoidal like function \( S(\delta) = B(\tilde{\nu}) \cos(2\pi \tilde{\nu} \delta) \) where \( B(\tilde{\nu}) \) is a dispersive intensity.
Figure 3.9: Schematic representation of an FTIR system including both transmission and reflection measurement modes.

Figure 3.10: (a) Single-sided, bi-directional interferogram and (b) Fourier transformed single beam spectrum of an Al₂O₃ substrate. Interferogram obtained using a resolution of 4 cm⁻¹ and 800 scans. Interferogram was phase corrected and spectrum calculated using the Mertz method. Forward and backward scans were transformed and phase corrected separately, and combined at the end using a mean function.

function, dependent on the source intensity, optical losses introduced by the instrument (e.g. beam-splitter efficiency, and detector and amplifier responsivity), and optical absorption from the sample and environment. By measuring the intensity for different retardation values, one can easily determine the wavelength and intensity of the signal by measuring the period and amplitude of the sinusoidal wave.
Real system limitations and workarounds

On a real (and more efficient) system, a continuum source with a broad emission range will be used instead. In this scenario, the interferogram will not be a sinusoidal function but will have a complex shape such as in Fig. 3.10.(a). This is the result of constructive and destructive interference between photons of different wavelengths, and it can be mathematically described as

\[ S_T(\delta) = 2 \int_0^\infty B(\tilde{\nu}) \cos(2\pi \tilde{\nu} \delta) d\tilde{\nu}. \]

The centerburst in Fig. 3.10.(a) (peak of highest intensity) is the result of all the waves being in-phase and constructively interfering. A Fourier transformation of \( S_T(\delta) \), with interferogram correctly centered and sampled uniformly, will lead us to obtain \( B(\tilde{\nu}) \) containing the spectral information of the interferogram signal. Since the interferogram is a function of retardation (not time), a unit conversion is required to obtain the appropriate wavenumber value \( \tilde{\nu} \). Considering a scanner speed \( V \), the retardation at time \( t \) can be obtained from \( \delta = 2Vt = 2V/f \), and so, \( f = 2V\tilde{\nu} \) where \( f \) is the frequency obtained from the Fourier transformation of \( S(\delta) \).

In practice, a commercial FTIR has only limited space for the moving mirror, meaning that data can only be obtained up to a certain point (\( \delta_{\text{max}} \)). This in turn affects the capability of the instrument to resolve photons of very similar wavelengths (resolution limited to \( \Delta\tilde{\nu} = \delta_{\text{max}}^{-1} \)), and can lead to artificial side lobes in the interferogram data. FTIR is also susceptible to phase effects due to high-frequency electronic noise and sampling effects, and for that reason, care has to be taken when computing the spectrum. In this work, the Mertz method was used to correct for artefacts introduced by the instrument [150]. This includes the application of an apodization function, to reduce the effect of the artificial slopes in the interferogram, a phase correction procedure, and a zerofilling function, which increases the number of points in the spectrum and smooths down the data. For the example in Fig. 3.10.(a), containing two centerbursts taken from the forward and backward mirror directions, each interferogram is transformed and phase corrected separately, and the final spectrum, shown in Fig. 3.10.(b), is obtained from a mean of the individual spectra. The result in Fig. 3.10.(b) was obtained from a transmission measurement through an Al\(_2\)O\(_3\) substrate. In practice, the globar
source has a much larger range with peak intensity around 8\(\mu\)m. However, due to the low Al\(_2\)O\(_3\) transmission for wavelengths larger than 5\(\mu\)m, there is a sharp cut-off above this range in the intensity plot.

**Background correction and FTIR microscope operation**

![Figure 3.11: Transmission of a ZnO film on an Al\(_2\)O\(_3\) substrate, normalized with respect to a background measurement of the substrate.](image)

Absorption of IR radiation can excite vibrational modes in a molecule, leading to a spectral signature dependent on the atomic composition. FTIR is therefore an invaluable tool for determining the chemical and molecular structure of various solids, liquids and gases. However, appropriate background correction is required to account for environment absorption effects. This situation is illustrated in Fig. 3.10.(b) corresponding to an FTIR measurement of an Al\(_2\)O\(_3\) substrate (no background correction), where absorption features are visible in regions of the spectrum corresponding to water vapour and carbon dioxide molecular vibrations. Correction against environment effects can be achieved by taking two separate measurements: (1) a measurement of the environment \(B_b(\tilde{\nu})\) including substrate materials and surrounding gases or liquids; and (2) a measurement of the sample
Chapter 3. Experimental and numerical methods

To minimize the changes in environment conditions between the two measurements, some systems are vacuum sealed, or have a special sample enclosure attached to a purge gas system (N₂ or dry air). This allows to maintain constant background conditions throughout the different measurements. Background correction is then performed by simply dividing the sample with the background spectral data \((B_s(\tilde{\nu})/B_b(\tilde{\nu}))\). A not so interesting result (but useful for later) is illustrated in Fig. 3.11, showing a transmission measurement of a ZnO film on an Al₂O₃ substrate, background corrected using the Al₂O₃ measurement in Fig. 3.11.(b).

Due to limitations in fabricating large samples all measurements were achieved using a commercial microscope (Bruker IRScopeII) attached to a FTIR (Bruker IFS66v/S), offering a visible field of view with 900µm diameter using a 15× IR objective. The system can be operated in transmission or reflection by simply activating a motorized mirror which deflects the IR beam from the spectrometer. In transmission, the collimated beam follows to a condenser allowing Köhler illumination of the sample, leading to an even light distribution and control over the light intensity by adjusting the condenser distance. In reflection, the light is focused using the 15× objective over a 900µm field of view diameter. Independently of the mode, transmitted or reflected light is captured by the same objective and goes through an iris aperture, which defines the sampling diameter and minimizes beam divergence effects. The iris diameter was selected based on the sample size, typically 100µm or higher. During reflection measurements, the microscope requires an optical component to deflect the beam from the interferometer towards the sample, while allowing reflected light to reach the detector. This is achieved by introducing a hole which is 50% covered by an IR reflective material. As a result there is a slight cost of reduced signal intensity at the detector, and larger noise when operating in reflection mode. A parabolic mirror (not shown) is then used to focus the light beam into the mercury cadmium telluride (MCT) detector, which is cooled down to liquid nitrogen temperatures to minimize thermally generated charge carriers.

With regards to the FTIR measurement of plasmonic devices, appropriate
background normalization is required based on the sample composition and the operation mode. In transmission, the sample measurement should be normalized with respect to the transparent substrate, while in reflection, one should use a highly reflective surface with similar properties to the plasmonic film. A sufficiently large number of scans (typically 800 to 1200), was used to improve the signal to noise ratio (SNR) in the measurement (SNR \( \propto \sqrt{\text{scans}} \)). With regards to resolution, this was usually set between 4 and 8 cm\(^{-1}\) depending on the spectral features of the plasmonic system. This is quite broad relative to the smallest resolution of the FTIR (0.25 cm\(^{-1}\)). For devices with a polarization dependent response, the incident light polarization, and/or polarization state of the transmitted/ reflected light, were modified by introducing polarizers in specific locations of the system, as highlighted in Fig. 3.9.

### 3.2.2 Light amplification by stimulated emission of radiation (laser)

A laser operates by exploiting the process of stimulated emission, where a photon stimulates a high energy atom to release its energy as another photon, with the same wavelength and phase. At thermodynamic equilibrium, most atoms in a material are in a low energy state reducing the chances of stimulated emission. To reverse this effect, an external light source is used to pump the gain medium and create a population inversion inside the lasing material. Once in an excited state, atoms release their energy very quickly by spontaneous emission of a photon (less than 10\(^{-8}\) s), making it difficult to reach other atoms for stimulating emission. To overcome this problem, lasers use materials containing energy levels with a longer lifetime (in the order of \( \mu s \) to \( ms \)), allowing them to build up a population inversion for stimulated emission- these levels are called metastable states.

Figure 3.12 shows the energy exchange of a four level laser system: (1) an external light source is used excite atoms from a lower \( E_0 \) to a higher \( E_1 \) energy level by photon absorption; (2) the excited atom will then release \( E_1- E_2 \) energy by spontaneous emission and populate the metastable state \( E_2 \); (3) some of the atoms in \( E_2 \) will spontaneous emit a photon with energy \( E_2- E_3 \) and because of
Chapter 3. Experimental and numerical methods

Figure 3.12: Energy levels in a four level laser system.

the population built up in this state, there is a higher probability of stimulated emission from other atoms in the metastable state; (4) atoms will then lose $E_3 - E_0$ energy by spontaneous emission reaching the ground state. The photons emitted in step (3) form a monochromatic and coherent light source. The first ruby laser [2] was actually a three-level system (no $E_3$ level), however, due to the high energy requirements to build up a population inversion, multiple energy level systems tend to be more efficient and practical to use.

A high-power commercial solid state laser is very often a complex optical system comprised of multiple oscillator and amplification stages. An oscillator in its simpler form consists of a gain medium with two mirrors at both ends, where a coherent light source is produced by bouncing back and forward photons generated by spontaneous and stimulated emission, triggering a chain reaction and a building-up of energy. The resonant wavelength of the oscillator is set by the gain medium and the optical properties of the oscillator components (mirrors, host material and diffraction gratings or prism). The release of energy can be achieved by: having a partially reflecting mirror, allowing transmission of light when a certain energy level is reached; through Q-switching, where the quality-factor of the resonant cavity is modulated to build up a large population inversion and then suddenly release all its energy in the form of a very short pulse; or through mode-locking, where ultrashort pulses of light are modulated inside the cavity by
Gain medium

Beam diameter

Output wavelengths tunable (680 to 1080 nm)

Repetition rate, \( f_p \) 80 MHz

Pulse duration, \( t_p \) 140 fs

Duty cycle, \( f_p t_p \) 1.12E-5

Pulse energy, \( E_p \)

\[
P_{\text{peak}} = \frac{E_p}{t_p} = 314 \text{ kW @ 800nm}
\]

\[
P_{\text{avg}} = E_p f_p = 3.5 \text{ W @ 800nm}
\]

Table 3.5: System specification for the Chameleon\textsuperscript{TM} Ultra II laser used in this work.

locking the phases of the oscillator modes, and some of its energy released every round trip. In contrast to an oscillator, the optical amplifier is a single pass device where a continuous or pulse light beam is amplified using a high gain medium by stimulated emission.

In this work, we have used a mode-locked solid state Ti:Al\textsubscript{2}O\textsubscript{3} laser for the measurement of nonlinear optical effects, with the specification presented in Tab. 3.5. The Chameleon\textsuperscript{TM} Ultra II uses Ti:Al\textsubscript{2}O\textsubscript{3} technology with a broad absorption and emission spectrum, Fig. 3.13.(b), as a result of vibrational energy transitions. This property allows Ti:Al\textsubscript{2}O\textsubscript{3} laser to have a wide wavelength range and to generate very short femto-second pulses at high repetition rates, resulting in a slightly larger average power than Nd:YAG systems. Pumping of the 500 nm absorption band is usually achieved with SHG from a Nd laser or argon-ion laser.
3.2.3 Optical parametric oscillator

An optical parametric oscillator (OPO) can be used to extend the wavelength range of the laser systems above. It relies on a three wave mixing process for down-conversion from a high frequency, high intensity laser beam ($\omega_{\text{pump}}$) to a low frequency signal and idler beams ($\omega_{\text{signal}} = \omega_{\text{idler}} - \omega_{\text{pump}}$). When an idler is not available, both signal and idler beams can be generated through parametric fluorescence in a nonlinear crystal, where a high energy photon is used to generate two lower energy photons by spontaneous emission [152]. Amplification of the idler and signal beams is achieved through difference-frequency generation ($\omega_{\text{signal}} = \omega_{\text{pump}} - \omega_{\text{idler}}$), where $\omega_{\text{signal}}$ stimulates generation of $\omega_{\text{idler}}$, and likewise $\omega_{\text{idler}}$ stimulates generation of $\omega_{\text{signal}}$ as the beams travel across the length of the nonlinear crystal. Similarly to a laser, one can introduce an optical cavity consisting of two highly reflective mirrors between the nonlinear crystal, to induce an oscillation of the signal beam. The tunability and oscillation of the $\omega_{\text{signal}}$ beam can be controlled via the cavity optics, or orientation or temperature of the nonlinear crystal.
In this work, a Chameleon™ Compact OPO-Vis OPO system has been used in conjunction with the laser described above. This has five operation modes, with power spectra and system diagram shown in Fig. 3.14.(b). The $\omega_{\text{pump}}$ corresponds to the laser line directly from the Chameleon™ Ultra II from which $2\omega_{\text{pump}}$ is obtained using SHG. The $\omega_{\text{idler}}$ and $\omega_{\text{signal}}$ are generated from an OPO, and the $2\omega_{\text{signal}}$ from another SHG process. Due to optical losses and inefficiencies in the nonlinear processes, there is a significant reduction in the average power on this system. Also, a compromise exists between the wavelength range of the OPO and the power efficiency of the system, since nonlinear processes are highly sensitive to the optical frequencies involved. In spite of these limitations, the OPOs can generate highly coherent light beams with a similar pulse duration and repetition rate as the pump laser. The measured linewidth of this systems is 33\text{nm} when operating with $\lambda_{\text{pump}}=800\text{nm}$ and $\lambda_{\text{signal}}=1393\text{nm}$, corresponding to an output signal of $\lambda_{\text{idler}}=1879\text{nm}$.
3.2.4 Monochromator

A monochromator is a powerful optical instrument that allows to decompose a light beam into its spectral components, and measure light intensity as a function of wavelength. A schematic representation of a spectrograph is shown in Fig. 3.15.(a), consisting of an entrance slit, optical grating and detector. The function of the slit is to control the amount of light that enters the monochromator and the minimum image size that reaches the detector, and is important in defining the resolution and throughput of the system. Light from the slit is then collimated towards a diffraction grating which spatially separates photons of different wavelengths. Grating properties such as groove frequency and blaze angle, are important in defining the amount and angle of dispersion, and also the resolution of the spectrometer. The wavelength coverage of a spectrometer, defined by the groove frequency, is inversely proportional to the resolution of the system, and therefore commercial spectrometers tend to provide a choice of gratings to accommodate for different experimental requirements. Of relevance to this work is the spectral efficiency of the low-resolution grating presented in Fig. 3.15.(b), with 150 lines/mm, blaze wavelength of 500 nm and 0.88 nm resolution.

Figure 3.15: (a) Schematic diagram of an unfolded Czerny-Turner monochromator and (b) spectral efficiency of a 150 lines/mm and 500 nm blaze wavelength grating. Adapted from Andor Shamrock 303i specification manual.

The diffracted light from the grating is then directed towards the detector, which typically consists of ∼10µm size sensors arranged over an area covering all angles
of dispersed light, and a charge coupled device (CCD), which is used to readout the photon generated charges. In this arrangement, the resolution of the spectrometer is limited by the number and pixel size in the detector. Another important parameter is the focal length between the grating and the detector, as this will dictate the spread of all the different wavelengths, and the final resolution of the system.

3.2.5 Optical detection systems

In Sec. 2.2, we have seen that owing to the very small nonlinear susceptibility of materials \( \chi^{(2)} \sim 10^{-12} \text{m/V} \) and \( \chi^{(3)} \sim 10^{-20} \text{m}^2/\text{V}^2 \), nonlinear optical effects are expected to be orders of magnitude smaller than their linear counterpart, despite the quadratic and cubic dependence of SHG and THG to the incident electric-field. The intensity of nonlinear effects can be greatly improved by combining the microscopic response from individual atoms through phase matching of the nonlinear waves (Sec. 2.2.1). However, for very thin nonlinear samples, on the order of hundreds of nanometers, or surface nonlinear effects, which are generated within a few atomic layers of a material’s surface, the resulting nonlinear signals are still significantly small and require not only strong field intensities but also highly sensitive detection systems.

The most common photodetectors are usually grouped into two classes: thermal detectors and photon/quantum detectors, with each divided into their own sub set of categories. The former use a class of materials capable of absorbing electromagnetic radiation and converting it to heat. The change in temperature causes a modification of the electrical properties of the material which is manifested by a change in thermoelectric voltage, resistance or pyroelectric voltage. Since their temperature is not influenced by the photon energy of the incident radiation but only the total power (number of photons per unit time), this type of detectors is characterized for having a linear response over a wide range of wavelengths (Fig. 3.16). These are however known for very slow time response and poor sensitivity (thermal 10\( \mu \text{W-100W} \) and pyroelectric 10\( \mu \text{J-10J} \)), and usually find applications in power or energy meters of intense light radiation such as lasers.
Figure 3.16: Spectral detectivity ($D^*$) of various visible and IR detectors. This is a measure of the signal-to-noise performance of a detector, and generally a large value means good detectivity. The MCT data was obtained from the manufacturers test report (InfraRed Associates, Inc. FTIR-7-1.0), the pyroelectric detector data from Hamamatsu [153], and the PMT and CCD data from Newport [154].

Given the simpler physical mechanism, thermal detectors are typically cheaper and easier to operate.

Regarding photon/quantum detectors, light is absorbed through interaction with atoms or free electrons in a material, resulting in a change of energy distribution and the generation of a voltage or current signal proportional to the number of photons. In contrast to thermal detectors, these have a wavelength dependent response since they rely on the band structure of materials to define the allowed electronic transitions. These are however associated with improved signal-to-noise performance and extremely fast response times, but limited by thermal generation of charge carriers, requiring the use of cryogenic cooling which makes these bulky, and more expensive. Of interest to this work are the mercury cadmium telluride (MCT), charge coupled device (CCD), and photon multiplier tube (PMT) detectors with specific detectivity as shown in Fig. 3.16. This quantity is inversely proportional to the noise equivalent power (a measure
of the radiant flux necessary to match the background noise of the detector), and it allows a comparison between different detection technologies by accounting for the sensing area (i.e. a larger sensor will generate more noise but capture more light).

Common sources of noise in a detector are shot noise, generation-recombination noise, Johnson noise, flicker noise, readout noise and microphonic noise. Different techniques exist to account for these limitations, including: cooling the detector to reduce the effect of Johnson noise; increase the integration time to reduce effects of short noise; and use of specialized amplifier stages and stable power sources to reduce microphonic noise. Depending on the technology there is always a fundamental limitation in the sensitivity of a detector, but the general rule is that light of the visible spectrum is much easier to detect than IR. PMTs typically provide the best performance of all detectors owing to their noise-free amplification mechanism, but given the limited availability of photocathode materials it is limited to ultraviolet and visible wavelengths. Alternatively, in the near-IR range one can use a multichannel detector based on a quantum well technology. This has the advantage of easy integration to spectroscopy and imaging systems, owing to the large photosensitive area, and offer high quantum efficiency. Finally, for the mid-IR range MCT is a commonly used material in detectors. The optical absorption of the film can be easily tuned based on the cadmium composition, allowing the band gap to be tailored over a range from 1\(\mu\text{m}\) to 30\(\mu\text{m}\) [155]. They offer the best quantum efficiency performance of all detectors in the mid-IR, and they are commonly found in FTIR systems.

3.2.6 Measurement of nonlinear frequency generation

As discussed in Sec. 2.2, nonlinear optical effects require large field intensities to be excited to a level within current detection capabilities. Light emission from a continuous non-coherent source is usually too small to trigger a measurable optical response, so FTIR cannot be used. Only lasers are capable of producing sufficiently large field intensities that can be used to excite nonlinear optical effects, thanks to the energy-storage capabilities of laser gain medium. Measurement of nonlinear
phenomena therefore requires specialized optical arrangements and equipment. We will now look into some of the most common optical setups for measurement of nonlinear effects in transmission and reflection modes.

Figure 3.17: Optical setup for nonlinear transmission measurements.

Figure 3.17 shows an experimental setup for measurement of nonlinear optical effects in the near-IR in transmission mode. The excitation source consists of a laser and OPO, as discussed in Sec. 3.2.2 and Sec. 3.2.3, allowing the generation of very short and high intensity pulses over a range of wavelengths. From the laser source, the collimated linearly polarized light beam then goes through two polarizers $P_1$ and $P_2$. The function of the polarizers are two fold: (1) to control the pulse energy incident on the sample, and (2) to rotate the polarization of the electric field incident on the sample. The principle of operation of these polarizers is described by the Maluss law, where the intensity of a linearly polarized wave incident on a polarizer $P_1$ with intensity $I_0$ is proportional to $I_1 = I_0 \cos^2 \theta_1$, where $\theta_1$ is the angle difference between the light polarization of the laser and the angle of the polarizer. Physically this means that the vector component of the incident beam perpendicular to the transmission plane of the polarizer will be absorbed, reducing the amplitude of the light and changing the polarization angle (with respect to the optical table). In this example, the polarizer $P_1$ is only used to control the pulse energy of the sample, where $\theta_1 = 0^\circ$ gives maximum transmission and $\theta_1 = 90^\circ$ the minimum transmission through the polarizer.

If we then combine two polarizers in series, the same principle applies. The output intensity $I_2$ is proportional to the light intensity from the first polarizer $I_1$ times a factor $\cos^2(\theta_1 - \theta_2)$ that absorbs the vector component perpendicular to the transmission plane of the polarizer $P_2$. Overall pulse intensity is then expressed as
Figure 3.18: Relative laser intensity for different $\theta_1$ and $\theta_2$ polarizer angles according to Malus’s Law. Angle $\theta_2$ can be used to control the polarization incident on the sample with some compromise on the intensity of the light. Red line highlights the optimum combination of $\theta_1$ and $\theta_2$ which gives the highest laser intensity.

\[
I_2 = I_1 \cos^2(\theta_1 - \theta_2) = I_0 \cos^2(\theta_1) \cos^2(\theta_1 - \theta_2),
\]

and we essentially use the polarizer P2 to control the polarization of the electric field incident on the sample at a cost of lost energy. This situation is illustrated in Fig. 3.18, where the factor given by the contour lines, is the relative output pulse energy from the two polarizers, in respect to $I_0$, as a function of $\theta_1$ and $\theta_2$. The red line highlights the optimum relative energy for different electric field polarizations $\theta_2$ incident on the sample by setting of $\theta_1 = \theta_2/2$.

The monochromatic collimated beam then goes through a plano-convex lens L1 allowing to focus the light onto the nanostructures with a beam waist of around $2\omega_0 = 100\mu m$. This is something that requires some optimization, given that larger intensities can be achieved with a smaller beam waist, $I(\omega) = P_{\text{peak}}/\pi\omega_0^2$, allowing maximization of the nonlinear processes (Eq. 2.27). To facilitate the alignment, the devices were fabricated over an area twice as large as the beam waist. In particular for mid-IR wavelengths and opaque substrates, where there is no visual clue in the location of the focal point, one must correctly align the samples to optimize the
linear and nonlinear signals. After the lens, a long pass filter $F_1$ is introduced to remove any unwanted nonlinearities from the optical system. The sample is then positioned in a manual stage for roll, tilt and $(x,y)$ control with micrometer precision, allowing adjustment of the incident electric field polarization $\psi$, off-normal incidence angle $\theta$ and focal point on the sample. Following the sample, a lens or objective $L_2$ is used for collimating the point source, containing the linear and generated nonlinear signals from the sample, and directed towards a polarizer $P_3$, which in combination with $P_1$ can be used to address specific tensor components of the linear and nonlinear polarizations. This situation can be better understood by expanding the SHG polarization function as follows

$$
\begin{bmatrix}
P_x(2\omega) \\
P_y(2\omega) \\
P_z(2\omega)
\end{bmatrix} =
\begin{bmatrix}
\chi^{(2)}_{xxx} & \chi^{(2)}_{xxy} & \chi^{(2)}_{xyx} & \chi^{(2)}_{xyy} & \chi^{(2)}_{xzz} & \chi^{(2)}_{xyz} \\
\chi^{(2)}_{yxx} & \chi^{(2)}_{yyx} & \chi^{(2)}_{yyz} & \chi^{(2)}_{yzx} & \chi^{(2)}_{yyy} & \chi^{(2)}_{yyz} \\
\chi^{(2)}_{zxx} & \chi^{(2)}_{zxy} & \chi^{(2)}_{zyx} & \chi^{(2)}_{zyy} & \chi^{(2)}_{zzx} & \chi^{(2)}_{zxy}
\end{bmatrix}
\begin{bmatrix}
E_x(\omega)^2 \\
E_y(\omega)^2 \\
E_z(\omega)^2 \\
2E_y(\omega)E_z(\omega) \\
2E_x(\omega)E_z(\omega) \\
2E_x(\omega)E_y(\omega)
\end{bmatrix}.
$$

(3.1)

Table 3.6: Tensor components as function of different polarizer-analyzer configurations for normal incidence. The $x$ and $y$ components are relative to the plane of incidence of the sample.

<table>
<thead>
<tr>
<th></th>
<th>$x-x$</th>
<th>$x-y$</th>
<th>$y-x$</th>
<th>$y-y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^{(2)}_{xxx}$</td>
<td>$\chi^{(2)}_{yxx}$</td>
<td>$\chi^{(2)}_{zxx}$</td>
<td>$\chi^{(2)}_{yyy}$</td>
<td></td>
</tr>
</tbody>
</table>

Taking the case were the sample is a pure noncentrosymmetric material, at normal incidence only $E_x(\omega)$ and/or $E_y(\omega)$ electric-field components contribute to the SHG polarization. Control of their individual intensity is achieved using the polarizer $P_2$. At off-normal incidence, the $E_z(\omega)$ component can be also activated to trigger the effect of other nonlinear tensor components. The effect from individual components can be analyzed using the polarizer $P_3$, which are highlighted in Tab. 3.6 for the case using normal incidence. When considering the response from plasmonic nanostructures, the local fields $E(\omega, \mathbf{r})$ will lead to the activation of multiple tensor components, as result of LSP resonances and variations in charge distribution.
across the sample. In the context of this work, we are mainly interested in the global effect of the plasmonic enhancement, rather than measuring individual tensor contributions. For that reason, the analyzer P3 has not actually been used in experiments. The purpose of this analysis is only to serve as an illustration on the probing mechanism of different tensor components of a nonlinear medium.

The final stage of the nonlinear optical setup consists of a lens L3 and short pass filter F2, which are used to focus the collimated beam into the sensing area of the detector, and remove any signal at the fundamental frequency. Typically, for measurement of nonlinear signals, the detector is composed of a monochromator with a PMT or CCD detector attached, which is chosen for maximum sensitivity in the wavelength range of interest.

Figure 3.19: Schematic diagram of a confocal optical microscope used in the surface SHG measurements in Chap. 5. In the current configuration the microscope is operating in reflection mode.

In Sec. 5, an alternative system has been used for reflection measurements of surface SHG effects from plasmonic metasurfaces. This is based on confocal laser
scanning microscopy, which is a 3D high resolution optical microscope primarily used in the life sciences. A simplified schematic diagram of this instrument is shown in Fig. 3.19. Collimated lasing excitation, with linear polarization and fixed wavelength, passes through an illuminating aperture and long pass filter, to remove nonlinearities from the laser, and is directed towards the sample using a dichroic mirror. Typically these systems operate in the visible range, and therefore, the tunability offered by a Ti:Sapphire laser is sufficient for most applications. Before reaching the sample, linearly polarized light is converted to CPL using a λ/4 waveplate, and is focused onto the sample using a high numerical aperture, forming a beam waist smaller than \(2\omega_0 = 1\mu m\). Fundamental and SHG light reflected from the sample then follows the reverse path towards the dichroic mirror. This transmits the SHG light (all other wavelengths are reflected) following to a short pass filter, which removes any unwanted frequencies from the signal. Before reaching a PMT detector, light passes through one of the main components of a confocal microscope, i.e. a confocal aperture or pinhole, which removes any light coming from out-of-focus planes allowing imaging of a single focal plane. An image of the sample is constructed by moving the sample stage. This technique is widely used in the life sciences for making three dimensional images of live cells using a fluorescent dye. In the context of this work, the advantage of a confocal microscope over a benchtop setup is the fact that they are usually optimized to work in the visible range, allowing large optical intensities, which are useful for studying weak surface nonlinear effects.

Ablation effects

During the nonlinear measurements, we have encountered some issues related with the localized heating of the samples. In metals, laser light is absorbed by free electrons in a quick thermalization process that lasts a few femtoseconds. During this process, the free electrons will diffuse and transfer their energy to the lattice via electron-phonon scattering, with a relaxation time of a few picoseconds, and as a result the metal can melt and undergo a phase transition leading to ablation [156]. This effect can occur at lower laser intensities in nanostructured metal surfaces.
This is due to the excitation of plasmonic resonances that lead to an inhomogeneous distribution of the free electrons and consequently the formation of hot spots [157].

Figure 3.20: (a,b) SEM micrograph of two plasmonic samples exposed to 125TW/m$^2$ and 320TW/m$^2$ intensities from a Nd:YAG pico second laser, and (c) FTIR response of sample (b) before and after nonlinear measurements. The samples consist of an array of slit apertures on a gold film filled with GaAs. The dimensions of the slits are as described in the inset of picture (c). The plasmonic response is characterized by a single LSP resonance at 8µm, when excited with an electric-field perpendicular to the longer arm of the slit.

In Fig. 3.20, we show a SEM micrograph of a slit aperture array filled with GaAs, after illumination at pulse intensities (a) 125TW/m$^2$ and (b) 320TW/m$^2$. Formation of hotspots due to local melting of the plasmonic metal can be clearly seen in the SEM micrographs. At increasing energy, this effect has a destructive impact on the plasmonic structure, which largely affects the optical response as shown by Fig. 3.20.(c). This effect can be explained by the local melting of the gold film, due to the large laser intensity, and a combination of hydrodynamic effects from the LSP resonances, phase transition changes and a rapid cool down of the metal. As a result, all these effects lead to the formation of nano gold droplets, and
evaporation of some of the metal from the sample. Similar effects have already been reported with other metasurfaces [157]. We found that by integrating a material with a higher melting temperature, we were able to achieve a better diffusion of the heat across the sample and minimize ablations effects. In Chap. 5, no ablation effects were observed in the plasmonic samples covered with a ZnO film fabricated on a sapphire substrate. The samples were however subject to a much lower beam intensity of \(1.1\text{TW/m}^2\) from a Ti:Al\(_2\)O\(_3\) femtosecond laser. In the measurement of devices in Fig. 3.20, a high energy Nd:YAG pico second laser was used instead.

### 3.3 Modelling methods

Since the foundation of classical electrodynamics, with the definition of the Maxwell’s famous equations [20], a range of analytical and numerical techniques have been developed to model light and matter interactions. Analytical methods very often provide accurate and elegant solutions to EM problems. However, these are usually confined to a very narrow set of problems, and may not be translated to real world applications without a significant increase in complexity. Alternatively, one can use numerical methods to find an approximate solution to more complex problems.

Numerical electromagnetic techniques are usually split in two categories: time-domain and frequency-domain methods. Time-domain techniques are extremely powerful in modelling the transient behaviour of EM fields, in particular problems involving a wide range of frequencies, while frequency-domain methods can only solve for a single frequency, but can naturally deal with problems involving oscillatory processes or periodic structures with greater accuracy. Another distinction is in the form of spatial discretization, which can be in the form of finite difference or finite element methods. The former divides the simulation space in discrete rectangular grid points, which works well for rectangular structures but becomes an issue when the model has a more complex shape. This can usually be resolved by increasing the number of grid points around critical areas. However, not all
finite difference methods allow a variable discretization of the simulation space, leading to large memory and computational requirements. Alternatively, the finite element method can be employed, which divides the simulation space into a variable set of subdomains (typically using a tetrahedral mesh), allowing a more accurate representation of the model, at a larger cost of complexity and processing times. An issue with methods involving spatial discretization is that for scattering problems, they can lead to artificial reflections at the boundary of the grid points or subdomains. Typically an artificial layer is employed in the model to absorb the radiation from these reflections.

In the present work periodic arranged metasurfaces have been investigated using a range of materials and excitation conditions. A frequency-domain method was therefore employed to model the optical response and near-field distribution at fundamental frequency from the metasurfaces. The particular method we have used is known as the rigorous coupled wave analysis (RCWA), through the commercial software Synopsys® DiffractMOD™ v. 2013.12-2. An overview of the RCWA operation will be given in the following section.

### 3.3.1 Rigorous coupled wave analysis

The RCWA is a semi-analytical method capable of computing rigorously the diffraction efficiency of 2D or 3D periodic structures, i.e. the relative power of each diffraction order, produced by the grating, to the incident signal. To describe the workings of this method, we will first start by considering the 2D periodic grating in Fig. 3.21. The same process can be implemented for 3D geometries (periodicity in two dimensions) but for simplicity we will consider the 2D grating case based on the formulation in Moharam and Gaylord [158].

The model is composed of three regions where the semi-infinite regions I and III are made of a homogeneous material with permittivities $\varepsilon_1$ and $\varepsilon_3$ respectively, and region II is composed of a grating with a spatially varying permittivity $\varepsilon_2(x, z)$ of period $\Lambda$ and thickness $d$. The grating consisting of a periodic distribution
Chapter 3. Experimental and numerical methods

Figure 3.21: Example geometry of a 2D grating (adapted from Moharam [159]).

of materials with permittivity $\varepsilon_1$ and $\varepsilon_3$, can be expanded by a Fourier series

$$\varepsilon_2(x, z) = \varepsilon_2(x + \Lambda, z) = \sum_p \varepsilon_{2,p}(z) \exp(i p K x)$$

where $p$ is the harmonic index, $K = \frac{2\pi}{\Lambda}$ the magnitude of the grating vector, and $\varepsilon_{2,p}(z)$ the Fourier component of the grating permittivity. For the case of a more complex structure with a $z$ variation of the permittivity, one can split the grating into $N$ slices and perform the analysis on each slice by matching boundary conditions on each side of the interface. In this example, we are defining a simpler case where there is no $z$-dependence.

In this exercise, the structure is excited with a monochromatic wave of wavelength $\lambda$, incidence angle $\theta$, and with an electric field $\mathbf{E}$ normal to the $(x,z)$ plane. The wave vector of the incident wave is described by

$$\mathbf{k}_i = k_0 \sqrt{\varepsilon_1} [\sin(\theta) \hat{x} + \cos(\theta) \hat{z}]$$

where $k_0 = \frac{2\pi}{\lambda}$. Light reaching the grating will form diffraction waves that will reflect or transmit light. In region I, the total field is described by the sum of the incident and reflected fields, as follows

$$E^I = E^I \exp(i \mathbf{k}_i \cdot \mathbf{r}) + \sum_n R_n \exp(-i \mathbf{k}_n^I \cdot \mathbf{r})$$

where $E^I$ is the amplitude of the incident field, $\mathbf{r}$ the position vector, $n$ the diffraction order index, and $R_n$ the amplitude of the $n$-th order diffracted wave with wave vector $\mathbf{k}_n^I$. Similarly, we can obtain an expression for the field inside the grating (region II).
which is given by

\[ E^{II} = \sum_n S_n(z) \exp(-i k_n^{II} \cdot r) \]  

(3.3)

where \( S_n(z) \) is the amplitude of the space harmonic fields, \( k_n^{II} = k_0 \sqrt{\varepsilon_2} [\sin(\theta) \hat{x} + \cos(\theta) \hat{z}] - K \hat{x} \) the wave vector for the diffracted waves, and \( \varepsilon_2 \) the average permittivity of the grating. Finally, the expression for the total transmitted field in region III is obtained from

\[ E^{III} = \sum_n T_n \exp(-i k_n^{III} \cdot (r - d \hat{z})) \]  

(3.4)

where \( T_n \) is the amplitude of the n-th order transmitted wave with wave vector \( k_n^{III} \).

The diffracted waves in regions I and III need to be phase matched with the space harmonic fields in region II, requiring the x-component of the wave vectors to be equal, i.e. \( k_n^I \cdot \hat{x} = k_n^{II} \cdot \hat{x} = k_n^{III} \cdot \hat{x} \). From this and knowing that \( k_0^2 = k_x^2 + k_z^2 \), the normalized wave vectors on each region are as follows

\[ k_n^I \cdot r = (k_0 \sqrt{\varepsilon_1} \sin(\theta) - nK) x - \sqrt{k_0^2 \varepsilon_1 - (k_0 \sqrt{\varepsilon_1} \sin(\theta) - nK)^2} z, \]  

(3.5)

\[ k_n^{II} \cdot r = (k_0 \sqrt{\varepsilon_1} \sin(\theta) - nK) x - \sqrt{k_0^2 \varepsilon_2 - k_0^2 \varepsilon_1 \sin^2(\theta)} z, \]  

(3.6)

\[ k_n^{III} \cdot r = (k_0 \sqrt{\varepsilon_1} \sin(\theta) - nK) x - \sqrt{k_0^2 \varepsilon_3 - (k_0 \sqrt{\varepsilon_1} \sin(\theta) - nK)^2} (z - d), \]  

(3.7)

where \( R_n, S_n \) and \( T_n \) are the only three unknowns to this problem. A solution is found by calculating the values of \( S(z) \), which is achieved by taking Eq. 3.3 and Eq. 3.6, and introducing these in the wave/Helmholtz equation for transverse electric (TE) polarization in region II, \( \nabla^2 E^{II} + k^2 \varepsilon_2(x, z) E^{II} = 0 \), leading us to

\[ \frac{d^2 S_n(z)}{dz^2} - 2 k_n^2 \frac{d S_n(z)}{dz} - \left[ (k_x^2 - nK)^2 + (k_z^2)^2 \right] S_n(z) - k_0^2 \sum_p \varepsilon_{2,p}(z) S_{n-p}(z) = 0, \]  

(3.8)

consisting of an infinite number of coupled second-order differential equations. In this formulation, the n-th space harmonic component in the grating is coupled with the other components through the \( (n - p) \) relation. The number of Fourier components \( p \) used in the definition of the grating permittivity will therefore affect the accuracy of the calculation (more of this in the next section). An infinite number of harmonic components is required to find an exact solution, but given the limited
computational resources, one has to truncate the number of terms to find an approximate result. In Synopsys® DiffractMOD™ this is achieved by defining the number of diffraction orders (harmonics) $N$ which correspond to $p = 2N + 1$ expansion terms.

A solution to the coupled wave equation can be found by using a state-variables method from linear systems analysis [160]. Once $S_i$ is determined, the amplitudes of the reflected and transmitted diffracted orders ($R_n$, $T_n$) are obtained by setting the boundary conditions between the regions (I+II) and (II+III).

**Gibbs phenomenon**

Being a Fourier based method, RCWA suffers from the Gibbs phenomenon leading to jump discontinuities of the calculated fields between material interfaces [161]. This feature is illustrated in Fig. 3.22 for two samplings of the near field distribution: (a,b) inside a dielectric material and (c,d) at the surface of a metallic nanostructure. For the dielectric case, where there is a small contrast between the relative permittivity of the different materials ($\varepsilon_{\text{ZnO}} = 3.65$ and $\varepsilon_{\text{air}} = 1.0$), we observe some harmonic content in the electric-field distribution, which can be minimized by increasing the number of Fourier terms in the model, see Fig. 3.22.(b). Whereas for the metallic surface, owing to the large contrast between interfaces ($\varepsilon_{\text{ZnO}} = 3.65$ and $\Re[\varepsilon_{\text{Au}}] = -142$), we observe large jump discontinuities of the calculated fields between different materials, Fig. 3.22.(d). On this type of structure, convergence of the far-field response (i.e. transmission, reflection and absorption) or near-field inside a dielectric, can be usually achieved using $N=8$ harmonics. However, care needs to be taken when analyzing the near-field distribution of interfaces with a large contrast, and for this we typically used a large number of harmonics, $N \geq 16$. 

85
Figure 3.22: Illustration of the Gibbs phenomenon on the near-field distribution at two different interfaces: (a,b) a dielectric medium and (c,d) at the surface of a metal with complex permittivity. The model consists of a gold cross-shaped metallic structure on a Al\textsubscript{2}O\textsubscript{3} substrate and covered with a ZnO dielectric medium. The simulation was performed using an electric-field polarization parallel to the vertical arm of the cross, and wavelength matching the LSP resonance of the vertical arm. The near field distributions were calculated using N=16 harmonics and represent a z-cut across a (left) dielectric interface and a (right) metallic interface. Images (b,d) correspond to a x-cut across the plane highlighted by the dashed lines in figures (a,c) for three different simulations using N=[8, 16, 24] harmonics.
Chapter 4

Tunable bulk nonlinear effects using complementary metamaterials

4.1 Introduction

Metallic nanostructures of dimensions below the wavelength of light show remarkable optical properties due to their ability of trapping light through excitation of localized surface plasmons (LSPs) (Sec. 2.1.3). These are characterized by strong electric-field localization and are of particular interest in nonlinear optics, where processes have a power dependence on the fundamental field (Sec. 2.2). Enhanced surface nonlinear effects from metallic nanoparticles have been extensively reported but with reduced effectiveness due to the small active region in the metal (Sec. 2.2.2). It is well established that the electromagnetic wave associated with a surface plasmon has a larger penetration depth in dielectrics than in metals, and therefore larger conversion efficiencies can be obtained by integrating a transparent material with bulk nonlinear properties as the dielectric medium [16, 162, 163, 164, 165, 166, 167, 168]. Another property of LSP resonances is the sensitivity to the polarization of incident light, allowing control of the field distribution around the nanostructures and the optical response of the devices (Sec. 2.1.4). Nonlinear optical effects are dictated by the symmetry properties of a
material, and as such are also sensitive to the polarization. Largely unexplored is how the interaction of tunable LSP resonances with the internal symmetries of bulk nonlinear materials can be harvested to further enhance and control the nonlinear processes. Also unknown is the effectiveness of the nonlinear process between complementary nanostructures, since it is known from Babinet’s principle that field transformations occur between geometries (Sec. 2.1.6). Answering some of these questions is highly desirable for the development of future integrated and quantum photonic devices, requiring precise control over photon-photon interactions.

Here we report on enhanced and optical polarization-assisted tuning of bulk second- and third- harmonic generation processes, using two samples of complementary asymmetric cross-shaped plasmonic nanostructures covered with a ZnO film characterized by large $\chi^{(2)}$ and $\chi^{(3)}$ nonlinear susceptibilities. ZnO is readily available using a range of deposition techniques [111]. However, it finds limited applications in nonlinear optics due to its small birefringence [114, 115]. Coupling with LSP resonances will therefore significantly enhance the nonlinear processes in ZnO films of thickness around tens of nanometres, without the need of phase matching. The optical tunability of the nonlinear processes is derived from the control over the orthogonal LSP resonances in the cross shaped nanostructure. Symmetric crosses have degenerate LSP resonances, and are therefore insensitive to changes in polarization [91, 92, 169]. However the spatial properties of the LSPs can be modified by introducing an asymmetry in the cross, allowing two resonant modes to cohabit with an intensity controllable by the incident electric-field polarization [90, 94, 169, 80]. Interestingly, as a result of the field transformations predicted by Babinet’s principle, our complementary structures show different multipole contributions to the nonlinear polarization. Our experimental and numerical findings will be presented through a side-by-side comparison of the different optical effects obtained from the two complementary samples.
4.2 Device description

For our studies we use two samples of complementary asymmetric cross-shaped metamaterials, associated with extraordinary optical transmission or reflection at LSP resonances, and with known applications in wavelength and polarization filtering [90, 169]. For clarity of interpretation we call “transmissive” the sample made of a continuous gold film containing cross-shaped apertures, and “reflective” the sample containing the complementary structure (i.e. isolated metallic crosses). Figures 4.1.(a) and 4.1.(b) show the periodic arrays of transmissive and reflective crosses (respectively), fabricated on a sapphire (Al$_2$O$_3$) substrate and covered with a ZnO film. Inset images in Fig. 4.1 show schematic cross-section of a unit-cell, coloured to highlight the different layered materials. Samples were fabricated on a 100×100 µm$^2$ area (approximately 17,000 unit cells) and measured using a beam of similar size, specified in further sections.

Figure 4.1: Helium focused ion beam micrograph of an array of (a) transmissive and (b) reflective cross samples. The mean dimensions of the crosses in both samples are $W_V=220$ nm, $W_H=70$ nm, $L_H=620$ nm, $L_V=380$ nm and $P=760$ nm, with labels specified in Fig. 4.2. The illustrations and material labels serve only for differentiating between plasmonic devices. The micrograph samples are not gold or sapphire terminated but have a ZnO film on top of the metallic nanostructures.

The complementary samples were designed to have similar LSP resonances
Figure 4.2: (a) top and (b) side view of an asymmetric cross geometry, showing main cross dimensions, in-plane electric field polarization angle $\psi$, k-vector angle of incidence $\theta$, and azimuthal angle $\varphi$.

in the near-IR. Assuming the validity of Babinet’s principle for metamaterials (Sec. 2.1.6), the average dimensions of the crosses were matched between samples, these being $W_V=220\,\text{nm}$, $W_H=70\,\text{nm}$, $L_H=620\,\text{nm}$, $L_V=380\,\text{nm}$ and $P=760\,\text{nm}$, where $W$ and $L$ represent the width and length of the horizontal and vertical arms, and $\Lambda$ the periodicity of the crosses, as illustrated in Fig. 4.2.(a). With regards to the average thickness of the films in the planar metamaterials, these are $5\,\text{nm}$ titanium, $30\,\text{nm}$ gold, and $75\,\text{nm}$ ZnO. An interesting aspect of this geometry is that independent control over the orthogonal LSPs can be achieved by varying the width and/or length of the cross arms independently [91, 94], allowing to access a whole range of different wavelengths. For the particular case of the transmissive sample, since our main resonant effect originates from LSPs, the propagating surface plasmon polaritons which originate from scattering effects on the periodic grating have small contribution to the optical response at normal incidence. With this in mind, the change in periodicity of the complementary arrays has very little effect on the optical response of the LSP modes, apart from changing the amplitude of the transmitted or reflected signals in proportion to the number of unit cells [170, 90]. It is to note however, that higher frequency SPP modes (not shown) can be affected in structure, amplitude and position by the periodicity of the array [90]. This is the result of the propagation nature
of the SPPs, and its effect can be predicted by the theory of Wood’s anomalies [171].

With regards to the choice of materials, our main nonlinear medium is ZnO which is a non-centrosymmetric material that exhibits both second-order $\chi^{(2)}$ and third-order $\chi^{(3)}$ nonlinear susceptibilities, as large as 0.32 pm/V and $185 \times 10^{-20}$ m$^2$/V$^2$ (respectively). It has an extremely high melting point at 1977°C and wide transparency band in the visible-IR spectrum, making it a desirable material for nonlinear optical applications requiring high lasing intensities. Another advantage is the variety of growth techniques, allowing ZnO to be deposited on a range of inexpensive substrates. A recent review on the nonlinear optical properties of ZnO films and nanostructures can be found in Larciprete et al. [111].

As a plasmonic material, we have chosen gold due to its good LSP quality factor in the IR range [30], chemical stability in normal environment conditions, and compatibility with our ZnO growth mechanism, which here is molecular beam epitaxy (MBE). The titanium film is for adhesion purposes only and has a small effect on the plasmonic response. The substrate material was chosen based on growth requirements and the desire for low levels of background nonlinearities in the measurement. Al$_2$O$_3$ is a common substrate in the MBE growth of ZnO films [111]. It is a centrosymmetric material, so that even-order nonlinear processes are forbidden within the electric-dipole approximation, and therefore no bulk SHG (Sec. 2.2.1) can be observed in Al$_2$O$_3$. The Al$_2$O$_3$ substrate can however produce THG, but since the highest reported value for $\chi^{(3)}$ is a few orders of magnitude lower than that of ZnO ($8 \times 10^{-23}$ m$^2$/V$^2$ [172]), the contribution from Al$_2$O$_3$ is neglected. This is also confirmed in experiment with a reference sample without the ZnO film.

4.3 Fabrication and ZnO film characterization

The fabrication process for the complementary samples is summarized in Fig. 4.3, and is based on processes (D) and (E) discussed in Sec. 3.1.7. It essentially uses EBL and argon-ion milling to pattern the nanostructures in the metallic films.
Using this approach, rather than a more conventional lift-off method, we are able to overcome the charging effects from the EBL when patterning on a non-conductive substrate (Sec. 3.1.7). As the substrate we have used (0001) c-plane oriented, 500 µm thick, single side polished Al₂O₃ from SurfaceNet GmbH, which was initially coated with a 5 nm titanium adhesion layer and 30 nm gold film using electron-beam evaporation at a base pressure of around 10⁻⁷ mbar.

Figure 4.3: Process steps for the fabrication of an array of asymmetric (a) cross apertures and (b) opaque crosses using EBL, argon-ion milling and MBE.

The final device structure will be dictated by the tone of electron-beam resist. For the transmissive sample, the gold film was coated with a positive resist consisting of 200 nm of MicroChem 950 PMMA. For the reflective sample, we have used two layers of resist consisting of 85 nm of Dow Corning® XR-1541 HSQ based resist, and an underlayer of 100 nm of 950 PMMA used to improve resist stripping later in the process. The nanostructures were then patterned onto an 100×100µm² area using a RAITH150 Two direct write EBL with a 30 kV beam voltage, 20 µm aperture, and respectively 330 and 297 µAs/cm² beam dosages for the transmissive and reflective samples. The transmissive sample was then developed in a (1:3) solution of MIBK:IPA for 60 s, and the reflective sample in a MF-26A developer also for 60 s. At this stage, the reflective sample requires an extra step to transfer the features from the HSQ onto the PMMA, which is achieved using O₂ plasma ashing for 9.5 min at 0.4 mbar and 200 W.

Pattern transfer from the resist to the metallic films was achieved using an SVS V6000 argon-ion milling system for 16 min with a base pressure below 10⁻⁶ mbar,
Chapter 4. Tunable bulk nonlinear effects using complementary metamaterials

Table 4.1: MBE growth times, rates and final thickness for each sample and interface.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Growth time (min)</th>
<th>Rates (nm/min)</th>
<th>Final thickness (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>gold</td>
<td>sapphire</td>
</tr>
<tr>
<td>transmissive</td>
<td>22</td>
<td>2.1</td>
<td>1.4</td>
</tr>
<tr>
<td>reflective</td>
<td>25</td>
<td>3.0</td>
<td>2.7</td>
</tr>
</tbody>
</table>

390 V acceleration voltage, 200 V beam voltage, 10 mA beam current and 5 sccm argon flow. The samples were then immersed in acetone overnight, and a few seconds ultrasonic helped strip-off the resist on top of the nanostructures.

The final stage of the fabrication process consists of a low temperature growth of ZnO using MBE. Before loading the samples, an oxygen-terminated surface was achieved through a 30 min oxygen plasma treatment at 0.4 mbar and 200 W. For the MBE growth of ZnO we have used an oxygen flow of 3 sccm, a zinc beam flux equivalent pressure of $2.2 \times 10^{-9}$ Torr, and an operational power of the oxygen plasma source of 300 W. For the transmissive sample, we found preferential growth of ZnO on top of the gold film, and agglomeration of ZnO molecules at the edges of the nanostructures. This effect is visible on the inset of Fig. 4.1.(a). We were able to improve nucleation and growth of ZnO inside the aperture by fixing the substrate temperature to $T_s=250^\circ$C on both samples. In spite of the lower process temperature, different growth rates were found between samples and material interfaces, with the reflective sample having the most uniform film of the two samples. Details on the growth times, rates and final thickness of the ZnO film for the different samples are shown in Tab. 4.1.

Figure 4.4.(a) shows the XRD patterns of the complementary samples, indicating that these are ZnO films grown with c-axis (0002) orientation, similarly to the substrate. A peak shift for the reflective sample suggests there is a compressive stress on the polycrystalline grains [173]. The broadening of the peak is associated with the formation of smaller polycrystalline structures of ZnO, also noticeable in Fig. 4.1. The pole figure measurements in Fig. 4.4.(b), reveal a six-fold symmetry of
Figure 4.4: (a) XRD measurement of transmissive and reflective samples showing peak at $2\theta/\omega=34.4^\circ$ corresponding to polycrystalline ZnO film with (0002) orientation. (b) Pole figure XRD measurement on both samples showing preferred orientation of the polycrystalline grains with respect to the nanostructures.

the films, confirming that the ZnO films are epitaxial. Differences in the XRD data suggest that the ZnO film on the transmissive sample has improved crystallinity (narrower $2\theta$ peak and sharper features on the pole data). We have attempted two growths on each sample and the results were reproducible. This result was not unexpected since most of the substrate interface of the reflective sample has the best crystallinity, coming from the (0001) c-plane oriented $\text{Al}_2\text{O}_3$ substrate, while for the transmissive sample most of the growth is on an amorphous gold film. We suspect that the difference in substrate temperature between samples might play a role in the alignment of the ZnO molecules has they condensate in the substrate (higher the temperature, longer the time to find the orientation of minimum energy). We were however unable to investigate this effect in more detail. Given the differences in crystallinity, it is therefore expected differences in the nonlinear susceptibilities between samples. Based on reported values for a range of process conditions and techniques, these variations are typically within the same order of magnitude [111]. Lastly, the polycrystallinity observed in the ZnO films (Fig. 4.1) is expected to contribute positively to the nonlinear effects. The nonlinear susceptibility is a material parameter highly dependent on the arrangement of atoms in the materials crystal structure. For the particular case of $\chi^{(2)}$, the added number of grain
boundaries and defects represent a break in the symmetry of the crystalline structure, and as such, larger values of the nonlinear susceptibility are expected [174, 175].

4.4 Numerical methods

In this section, we present a method for calculating the radiated linear, SHG and THG power through a periodic grating containing a material with bulk nonlinear properties. Typically nonlinear events in materials with bulk nonlinear properties are associated with electric dipole interactions of atomic excitations within the material. Higher-order multipoles can also play a role when the material is structured to dimensions below the wavelength of light [176], and/or through integration with certain resonant plasmonic nanostructures [131]. Here we use a multipolar analysis of the linear and nonlinear polarizations to model the optical response of the devices in Fig. 4.1.

ZnO has a wurtzite crystal structure which lacks inversion symmetry, and being of the space group 6mm, we have taken the following nonvanishing and independent tensor relations: \( \chi^{(2)}_{33} = 0.32 \text{ pm/V} \), \( \chi^{(2)}_{15} = \chi^{(2)}_{24} = 0.16 \text{ pm/V} \) and \( \chi^{(2)}_{31} = \chi^{(2)}_{32} = 0.16 \text{ pm/V} \) for the second-order nonlinear susceptibility; and \( \chi^{(3)}_{33} = 185 \times 10^{-20} \text{ m}^2/\text{V}^2 \), \( \chi^{(3)}_{11} = \chi^{(3)}_{12} = 1.99 \times 10^{-22} \text{ m}^2/\text{V}^2 \) and \( \chi^{(3)}_{18} = \chi^{(3)}_{29} = \chi^{(3)}_{11}/3 \) for the third-order nonlinear susceptibility [107, 177]. In our model, we only include polarization terms within the ZnO material, and thus ignore any surface SHG from the gold film, or THG from the Al₂O₃ substrate, gold or air as these are orders of magnitude lower than of bulk nonlinear processes in the ZnO [178, 172, 179]. Considering these simplifications and taking the case of a plane wave excitation with frequency \( \omega \), the linear \( \mathbf{P}(\omega, \mathbf{r}) \), second-harmonic \( \mathbf{P}(2\omega, \mathbf{r}) \) and third-harmonic
Chapter 4. Tunable bulk nonlinear effects using complementary metamaterials

\( \mathbf{P}(3\omega, \mathbf{r}) \) polarizations induced in the ZnO film can be obtained from [107]

\[
\begin{align*}
\mathbf{P}(\omega, \mathbf{r}) &= \varepsilon_0 (\varepsilon_r - 1) \mathbf{E}_i(\omega, \mathbf{r}) \\
\mathbf{P}(2\omega, \mathbf{r}) &= \varepsilon_0 \begin{bmatrix}
2\chi_{15}^{(2)} E_x(\omega, \mathbf{r})E_x(\omega, \mathbf{r}) \\
2\chi_{15}^{(2)} E_y(\omega, \mathbf{r})E_x(\omega, \mathbf{r}) \\
\chi_{31}^{(2)} (E_x^2(\omega, \mathbf{r}) + E_y^2(\omega, \mathbf{r})) + \chi_{33}^{(2)} E_z^2(\omega, \mathbf{r})
\end{bmatrix} \\
\mathbf{P}(3\omega, \mathbf{r}) &= \varepsilon_0 \begin{bmatrix}
\chi_{11}^{(3)} (E_x^3(\omega, \mathbf{r}) + E_x(\omega, \mathbf{r})E_y(\omega, \mathbf{r})^2) \\
\chi_{11}^{(3)} (E_y^3(\omega, \mathbf{r}) + E_x(\omega, \mathbf{r})^2E_y(\omega, \mathbf{r})) \\
\chi_{33}^{(3)} E_z^3(\omega, \mathbf{r})
\end{bmatrix}
\end{align*}
\]

where \( \varepsilon_0 \) is the vacuum permittivity, \( \varepsilon_r = 3.65 \) the relative permittivity of ZnO, and \( \mathbf{E}_i \) the i-component of the electric field inside the ZnO at position \( \mathbf{r} \). For the calculation of the fundamental electric-field \( \mathbf{E}(\omega, \mathbf{r}) \), we have used a commercial implementation of the RCWA method (Synopsys DiffractMOD v2013.12-2), as described in Sec. 3.3.1. The model was built using the experimental parameters as defined in Sec. 4.2 for the cross dimensions. For simplicity we have used a step-function in the definition of the ZnO film at the edges of the cross structures with a fixed thickness of 75 nm. We have also given fixed values for the refractive indices of \( n(\text{air})=1 \), \( n(\text{Al}_2\text{O}_3)=1.73 \) and \( n(\text{ZnO})=1.91 \), and have considered dispersive properties for the gold and titanium films using a dielectric function based on the Lorentz-Drude Model [31]. Numerical convergence on RCWA was obtained using \( N=8 \) harmonics for each transverse component, leading to \((2N+1)^2=289\) diffraction orders.

For the calculation of the lattice powers, we have used a multipolar expansion, limited to second-order terms only, namely the electric dipole \( \mathbf{p}(\omega_g) \), magnetic dipole \( \mathbf{m}(\omega_g) \), and electric-quadrupole \( Q_{ij}(\omega_g) \). The main set of multipolar equations as function of the polarization are given [34]

\[
\begin{align*}
\mathbf{p}(\omega_g) &= \int \mathbf{P}(\omega_g, \mathbf{r})d\mathbf{r} \\
\mathbf{m}(\omega_g) &= -i \frac{\omega_g}{2} \int \mathbf{r} \times \mathbf{P}(\omega_g, \mathbf{r})d\mathbf{r} \\
Q_{ij}(\omega_g) &= \int \{ 3 [r_i P_j(\omega_g, \mathbf{r}) + r_j P_i(\omega_g, \mathbf{r})] - 2\delta_{ij} \mathbf{r} \cdot \mathbf{P}(\omega_g, \mathbf{r}) \} d\mathbf{r}
\end{align*}
\]
where $\delta_{ij}$ is the Kronecker delta, and $\omega_g$ is the polarization frequency, i.e. $\omega$ for linear, $2\omega$ for SHG and $3\omega$ for THG polarizations. The total power $S$ emitted by the periodic array can then be calculated as \[ S(\omega_g) = \frac{Z_0 \omega_g^2}{8 A_0 \cos^2(\theta)} \left| \frac{(n \times p(\omega_g)) \times n}{S_{ed}(\omega_g)} - \frac{1}{c} \frac[n \times m(\omega_g)]{S_{md}(\omega_g)} - i \frac{k}{6} \left\{ [n \times (Q_{ij}(\omega_g) \cdot n)] \times n \right\} \right|^2 \tag{4.3} \]

where $Z_0$ is the impedance of free space, $A_0 = \Lambda^2$ the area of the unit cell obtained from the device period $\Lambda$, $n = \{\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta\}$ the plane wave direction vector, and $\theta$ and $\varphi$ the incidence and azimuthal angles as defined in Fig. 4.1.

4.5 Linear optical characterization

4.5.1 Transmission and reflection measurements on complementary samples

We have used a Bruker IFS66v/S FTIR spectrometer in combination with a Bruker IRscope II microscope to characterize the linear optical response of the samples in Fig. 4.1. An interferogram was formed in the spectrometer using a globar source of unpolarized IR radiation (7500-30 cm$^{-1}$) and a KBr beamsplitter (7500-370 cm$^{-1}$), and focused onto the sample using a 15$\times$ IR objective with 24mm working distance and a sampling diameter of 33$\mu$m. The transmitted and reflected signals from the sample were then detected using an MCT detector. Further information on the principles of operation of an FTIR can be found in Sec. 3.2.1.

Figure 4.5 shows (a,b) transmission and (c,d) reflection measurements for the (a,c) transmissive and (b,d) reflective samples using different electric-field polarizations $\psi$ and at normal incidence $\theta = 0^\circ$, as defined in Fig. 4.2.(b). The FTIR measurements were normalized with respect to a background of Al$_2$O$_3$ and ZnO in transmission, and of Al$_2$O$_3$, titanium, gold and ZnO in reflection, with both measurements being taken from unstructured areas of the sample. As seen in Fig. 4.5, the samples are characterized by having two modes, $\lambda_1$ and $\lambda_2$, that
Figure 4.5: (a,b) transmission and (c,d) reflection measurements for the (a,c) transmissive and (b,d) reflective samples using a FTIR. Measurement is achieved at normal incidence $\theta = 0^\circ$ with the electric field polarisation angle varying from $\psi = 0^\circ$ (black) to $\psi = 90^\circ$ (purple) with a $10^\circ$ step. Resolution $4\text{ cm}^{-1}$ and 800 scans used in both background and sample measurements. Dashed vertical lines highlight position of LSP resonances.

are associated with the coupling of incident light with different LSPs on each arm of the cross. Control over the intensity of these modes is achieved through change of the incident polarization $\psi$. Because the dimensions of the asymmetric crosses are similar in both samples, the modes are positioned around similar wavelengths, $\lambda_{T1}^{T} = 1888\text{ nm}$ and $\lambda_{T2}^{T} = 2667\text{ nm}$ for the transmissive sample, and $\lambda_{R1}^{R} = 1816\text{ nm}$ and $\lambda_{R2}^{R} = 2706\text{ nm}$ for the reflective sample. Discrepancies in the position of the modes are due to slight differences in the dimensions of the fabricated structures. The differences in amplitude between peaks in Fig. 4.5.(a) and 4.5.(d) are due to higher absorption losses in the transmissive sample because of the larger volume of gold.

We now introduce some numerical results regarding the linear response of
Chapter 4. Tunable bulk nonlinear effects using complementary metamaterials

the asymmetric cross-shaped devices. Figures 4.6.(a) and 4.6.(b) show the total power radiated from a lattice of multipoles calculated from the linear polarization (Eq. 4.1a) inside the ZnO film, for the transmissive and reflective samples respectively. These results show a good agreement with the transmission data obtained from the FTIR measurement in Fig. 4.5.(a) and 4.5.(b). Although the numerical model considers a very uniform ZnO film, which differs from the MBE film, the agreement in the linear results was to be expected since the ZnO is transparent in the wavelength of interest. Only small variations in amplitude are to be expected due to differences in transmission through the ZnO films (Fig. 3.11). This is because in the model a fixed value of the refractive index (non-dispersive) was used for ZnO.

An extra layer of information can be obtained in Fig. 4.6.(c-h) with regards to the most dominant multipole contributions. Independently of the incident polarization $\psi$, the electric dipole $S_{ed}$ is two orders of magnitude larger than both the magnetic dipole $S_{md}$ and the electric-quadrupole $S_{eq}$. For polarizations in the range $0^\circ < \psi < 90^\circ$, where the two LSP modes on each arm of the crosses can be excited simultaneously with varying intensity, these multipole contributions differ in magnitude, i.e. $S_{md} \neq S_{eq}$. In this interval, the electric-quadrupole is stronger for the transmissive sample, Fig. 4.7.(a), whereas the reflective samples has an equal magnetic dipole and electric-quadrupole contribution, Fig. 4.7.(b). Also, these are one order of magnitude higher in the transmissive sample (see insets in Fig. 4.7).

The nonlinear optical response from the metasurfaces are not only function of the nonlinear polarizations (Eq. 4.1b and Eq. 4.1c) but also the linear polarization at the SHG and THG wavelengths (Eq. 4.1a). In Fig. 4.8, the linear response of the complementary metasurfaces is shown for SHG and THG wavelengths, which will become useful in Sec. 4.6.5 when looking at the wavelength dependence of the nonlinear processes.
Figure 4.6: Numerically calculated (total) transmitted power from the (a) transmissive and (b) reflective samples at fundamental wavelength, normal incidence $\theta = 0^\circ$ and varying incident polarization $\psi=0^\circ$ to $90^\circ$ with a $10^\circ$ step. Dashed vertical lines highlight position of LSP resonances. Plots (c,d,e) and (f,g,h) show the electric dipole, magnetic dipole and electric quadrupole contributions, for the transmissive and reflective samples respectively. Individual multipole contributions normalized with respect to the peak power $S(\omega)$. 

100
Figure 4.7: Numerically calculated transmitted linear power from individual and total multipole contributions for the (a) transmissive and (b) reflective samples at resonant wavelengths $\lambda_T^1 = 1866$ nm and $\lambda_R^1 = 1828$ nm respectively. Source was set at normal incidence $\theta = 0^\circ$ and data normalized with respect to maximum total power intensity.

Figure 4.8: Numerically calculated (total) linear transmitted power from the (a) transmissive and (b) reflective samples for wavelengths corresponding to the SHG and THG of the fundamental range in Fig. 4.6. Simulation was performed at normal incidence $\theta = 0^\circ$ and using incident polarizations $\psi = [0^\circ, 90^\circ]$. Results were normalized with respect to the maximum total power in Fig. 4.6.
4.5.2 Babinet’s principle and field enhancement

The results from the FTIR measurements and linear numerical calculations are a clear example of the applicability of Babinet’s principle in the study of complementary plasmonic nanostructures. At resonance the devices are associated with enhanced-transmission or reflection in the case of an aperture (transmissive sample) or an opaque object (reflective sample) respectively, if the polarization of the incident plane-wave is orthogonal between samples. To understand this, let us consider a plane wave with EM-fields $E_i$, $H_i$ incident on the cross samples made of PEC films. According to Babinet’s principle (Sec. 2.1.6), the sum of the total EM-fields away from the complementary samples is related to the incident field of the original source by $E_t^o - uH_t^c = E_i^i$ and $H_t^o + E_t^c/u = H_i^i$, where the subscripts (o,c) refer to the fields from the original (transmissive) and complementary (reflective) objects, and $u$ is a conversion factor of units $\Omega$. This can be further simplified if we consider that at resonance ($\lambda_1$ or $\lambda_2$) there is a strong localization of the fields leading to $E_i \ll (E_t^o, u|H_t^i|)$. In this case, the incident field can be neglected and a direct relationship between the total EM fields on the transmissive and reflective samples can be obtained from $E_t^o = uH_t^c$ and $H_t^o = -E_t^c/u$. To visualize this, we have plotted in Fig. 4.9 the simulated electric- (top) and magnetic- (bottom) fields on the gold-ZnO interface of the (a-d) transmissive and (e-h) reflective samples. For a given resonance ($\lambda_1$ or $\lambda_2$), there is an interchange of the electric- and magnetic-field distributions between complementary samples, when the structures are excited with complementary (orthogonal) polarizations. As an example, the electric-field profile of the transmissive sample in Fig. 4.9.(a) matches the magnetic-field profile of the reflective sample in Fig. 4.9.(h), where polarization is rotated by 90°. Similar result has been already observed with split ring resonators [103]. Slight differences are the result of losses in the metallic film, in particular on the transmissive sample where there is a larger volume of gold.

Also of importance for this work is the electric-field distribution within the dielectric medium, in our case the nonlinear ZnO material. In Fig. 4.10 a cross section of the absolute linear polarization $|P(\omega, \mathbf{r})|$ is presented for the (a,b)
Figure 4.9: Simulated normalized EM-fields across the (a-d) transmissive and (e-g) reflective crosses when excited at plasmonic resonances $\lambda_1$ and $\lambda_2$ with $\psi = 0^\circ$ and $\psi = 90^\circ$ polarizations (respectively). Top and bottom rows show electric- and magnetic-field profiles (respectively) for the xy-plane at the gold-ZnO interface ($z=55$ nm in Fig. 4.10). The black arrows give the corresponding $|E|$ and $|B|$ field directions.

Figure 4.10: Cross section of the linear polarization $|P(\omega,r)|$ across the (a-b) transmissive and (c-d) reflective crosses when excited at plasmonic resonances $\lambda_1$ and $\lambda_2$ with $\psi = 0^\circ$ and $\psi = 90^\circ$ polarizations respectively. The black arrows give the corresponding $|P(\omega)|$ linear polarization direction.

transmissive and (c,d) reflective samples, using $\psi=0^\circ$ and $90^\circ$ incident polarizations. Only spatial points inside the ZnO nonlinear material were considered in the calculation of the linear polarization; areas containing Al$_2$O$_3$, titanium, gold or air were colored in blue, silver, orange or white (respectively). The linear polarization $P(\omega,r)$ has a linear relationship with the fundamental electric-field $E(\omega,r)$, as defined by Eq. 4.1a, and therefore the amplitude and direction of
Table 4.2: Average linear polarization density $P(\omega, r)$, in units of Cm$^{-2}$, inside the ZnO film on the complementary samples. The values correspond to different LSP excitations.

| sample      | $\lambda$ (nm) | $\psi$ (deg) | $|P_x(\omega)|$  | $|P_y(\omega)|$  | $|P_z(\omega)|$  | $|P(\omega)|$  |
|-------------|----------------|--------------|----------------|----------------|----------------|----------------|
| transmissive| 1866           | 0°           | 2.51E-11        | 6.95E-12       | 9.69E-12       | 3.14E-11       |
|             | 2588           | 90°          | 8.61E-12        | 2.66E-11       | 1.31E-11       | 3.47E-11       |
| reflective  | 1828           | 90°          | 1.43E-11        | 2.60E-11       | 1.00E-11       | 3.63E-11       |
|             | 2740           | 0°           | 2.02E-11        | 1.61E-11       | 8.70E-12       | 3.23E-11       |

The different components of $P(\omega, r)$ are expected to mirror the electric field $E(\omega, r)$. In Tab. 4.2, the average absolute value and individual components of the linear polarization density is presented for the main LSP resonances of the complementary samples. These data shows that the largest electric polarization occurs for excitation of a LSP mode in the vertical arm of the cross, which has the largest surface area. Another observation is that the main contribution is dependent on the incident electric-field polarization, these being $|P_x(\omega)|$ for $\psi=0^\circ$ and $|P_y(\omega)|$ for $\psi=90^\circ$. Interestingly, the contribution from $|P_z(\omega)|$ varies between complementary samples. For the transmissive sample, this is usually the second largest, whereas for the reflective sample this is the smallest of the three components.

The reason for the differences in the electric-field distribution between complementary structures is strongly associated with the resonant mechanism. For the transmissive sample, the LSP resonance is associated with a waveguide mode supported by the apertures, thus resulting in a strong field localization in these areas, whereas LSP resonances in the reflective sample, are the result of geometrical confinement of electrons in the nanostructure, leading to the formation of electric-field hot-spots at the edges of the cross arms. Recalling that the nonlinear processes have a power dependence on the fundamental electric-field (Sec. 2.2), it is expected that the nonlinear response between complementary samples will have some variations. In spite of the different field distributions, the strong electric-field confinement inside regions containing ZnO will provide an
Chapter 4. Tunable bulk nonlinear effects using complementary metamaterials

enhancement of the nonlinear processes. From the cross-section in Fig. 4.10, most of the enhancement originates from regions within the thickness of the metallic films.

4.5.3 Effect of the ZnO film on the LSP resonances

To understand the optical effect of adding the ZnO film on the samples, we have performed an FTIR measurement before the MBE growth. A reflection measurement on the reflective sample is presented in Fig. 4.11, showing that the resonant peaks are shifted by $\Delta \lambda^R_1 = 195$ nm and $\Delta \lambda^R_2 = 504$ nm in comparison to the results from the same sample with the ZnO film shown in Fig. 4.5.(d). From the Rayleigh approximation of scattering effects in a metallic sphere (Sec. 2.1.4), it is known that LSP resonances are sensitive to the relative permittivity of the dielectric medium, and it is therefore expected to see a redshift in the LSP resonances as we introduce the ZnO film ($\varepsilon_{\text{ZnO}}=3.65 > \varepsilon_{\text{air}}=1$). In terms of amplitude, the reflection results cannot be directly compared, as they are not corrected using the same background, i.e. optical losses from the ZnO are only considered on the measurement of the final samples (Fig. 4.5). As shown in Fig. 3.11 of Sec. 3.2.1, ZnO has 50% transparency in this wavelength range, and as such, some optical losses are expected by introducing the film.
Figure 4.11: Reflection measurement at normal incidence $\theta=0^\circ$ on reflective sample before ZnO growth. Resolution 4cm$^{-1}$ and 800 scans used in the measurement.

4.5.4 Effect of scaling the unit cell on the LSP resonances

An interesting aspect of LSP resonances is that they originate from the geometrical confinement of conduction electrons in the metal, and therefore are highly sensitive to the shape and size of the nanostructures. When fabricating the samples in Fig. 4.1, we also produced a series of devices of different sizes. Figure 4.12 shows the position of the LSP resonances ($\lambda_1^R$ and $\lambda_2^R$) for 4 different devices, based on the reflective design (without the ZnO film). From these results we observe a red-shift of the resonances as we scale the device dimensions. Fine tuning of a single resonance can be achieved through changing, individually, the width and/or length of one arm of the cross. This has been already reported elsewhere [94], and it serves as an illustration of the versatility of the asymmetric cross design.
4.5.5 Output polarization state

The FTIR measurements discussed above, do not contain any information regarding the polarization state of the transmitted or reflected light. Although the incident light is linearly polarized, we cannot assume that the polarization state will be maintained once it interacts with the devices. Information regarding the output polarization can be recorded by introducing a polarizer (aka analyzer) in the optical path just before the detector (Fig. 3.9). Figure 4.13, shows a set of FTIR measurements of the (top) transmissive and (bottom) reflective samples for seven different incident polarizations $\psi$, and 29 analyzer angles (from $-90^\circ$ to $90^\circ$ with a step of $6.43^\circ$). A background measurement was taken for every polarizer-analyzer combination, using unstructured areas of the samples (similarly to above), and used in the normalization of the sample data.

The data in Fig. 4.13 gives an indication of the polarization of the outgoing
(top) transmitted or (bottom) reflected light, as linearly polarized light interacts with the transmissive and reflective samples (respectively). If the polarization was to be preserved, we would expect to see regions of high transmission/reflection at identical polarizer-analyzer angles. However, this is not the case and we observe a spread of the polarization around LSP resonances (highlighted by the dashed lines) which suggests that the complementary plasmonic nanostructures change the polarization state of light. A complete description of the polarization state would require deriving the Stoke Parameters [34], which has information on the intensity, polarization, ellipticity and handedness of light. With our current FTIR setup, we are unable to derive information on the ellipticity and handedness, as this would require introducing extra optical components which our system cannot support [180].
Figure 4.13: Transmission (top) and reflection (bottom) measurements using a polarizer and analyzer configuration. Each column represents a measurement of one of seven incident polarizations $\psi$, for a range of wavelengths and 29 analyzer angles. The analyzer angle is defined with respect to the x-axis of the nanostructures. Top and bottom rows correspond to data for a transmissive and reflective sample (respectively). FTIR measurements are slightly noisier when operating in reflection, and for that reason, we have used a reflective sample with a larger unit cell to improve SNR. A factor of $1.875 \times$ was used on all unit cell dimensions of the reflective sample, to red-shift the LSP resonances towards a measurable range. Dashed white lines highlight resonant peaks, i.e. $\lambda^T_1=1888$ nm and $\lambda^T_2=2667$ nm for the transmissive sample, and $\lambda^R_2=2960$ nm and $\lambda^R_2=3770$ nm for the reflective sample. Resolution of 16cm$^{-1}$ and 200 scans were used in the background and sample measurement.
Chapter 4. Tunable bulk nonlinear effects using complementary metamaterials

4.6 Nonlinear optical characterization

In this section, we will look into bulk SHG and THG effects from the complementary nanostructures in Fig. 4.1. We will analyse the nonlinear optical effects, both experimentally and numerically, as a function of the laser power, incident electric-field polarization $\psi$, and pump wavelength $\lambda$.

4.6.1 Experimental methods

The optical setup used for the nonlinear measurements is summarized in Fig. 4.14, and it consists of a Coherent Chameleon Compact OPO pumped by a Coherent Chameleon Ultra II 140 fs, 80 MHz laser with a spectral range between 1750-4000 nm (idler mode) and 275-50 mW (25-5 kW) average (peak) power. The linearly-polarized collimated beam passes through a B-Halle achromatic quartz half-wavelength retarder, and a Thorlabs DGL10 double Glan-Taylor polarizer allowing control over the power and electric-field polarization $\psi$ incident on the sample (respectively). Any nonlinearities from the optical system that could affect the measurement, are removed using a combination of long pass optical filters with cut-offs at 1000 nm or 1500 nm (Thorlabs FGL1000 or FEL1500). Light is focused on the nanostructures, from the unpatterned side of the sample, using a Thorlabs LA1509 plano-convex lens with a 100 mm focal length and beam diameter around 100 $\mu$m. The nonlinear signals have a power dependence to the incident intensity, and as such, the nanostructures are positioned at the focal point to obtain the best signal. The sample is fixed onto a XYZ translation stage with rotational control. Light transmitted from the nanostructures, containing both fundamental and nonlinear signals, is then imaged using a Nikon CFI LU Plan Fluor EPI P 20× objective with 0.45 numerical apertures and 4.5 mm working distance, and coupled into an Andor Shamrock 303i Czerny-Turner spectrograph using a Thorlabs AC254-075-A-ML achromatic doublet lens with a 75 mm focal length. The spectrometer was configured to measure both SHG and THG signals simultaneously using a 150 lines/mm and 500 nm blaze grating with 0.88 nm resolution, and the signals measured using an Andor iDus 420 BR-DD air-cooled Si-CCD detector.
Chapter 4. Tunable bulk nonlinear effects using complementary metamaterials

with a limited spectral range from 300 to 1000 nm for a quantum efficiency over 20%.

Figure 4.14: Schematics of the experimental setup for measuring bulk nonlinear harmonic generation from the complementary plasmonic samples. A half-retarder and polarizer are used to control the power and incident electric field polarisation from a femto-second laser, spectrally tunable using an optical parametric oscillator system. The collimated beam is then focused onto the nanostructures, from the unstructured side of the sample, and a long pass filter is used to remove unwanted nonlinearities from the optical system. The transmitted light, containing both fundamental and harmonic generated signals, is then captured by an objective and coupled into a spectograph, which spatially separates photons of different energies and records these by using an air-cooled Si-CCD detector.

4.6.2 Power dependence

Figure 4.15 shows a power dependence measurement of the (green) SHG and (blue) THG signals at normal incidence $\theta=0^\circ$, with pump wavelengths and polarizations $\lambda_T = 1884 \text{ nm}$ and $\psi = 0^\circ$ for the transmissive sample, and $\lambda_T = 1804 \text{ nm}$ and $\psi = 90^\circ$ for the reflective sample. First and second rows are the SHG and THG data as measured from the spectrometer for the (left) transmissive and (right) reflective samples. The power from the laser was fixed to 105 mW and the half-wave plate was used to control the incident power, measured using a power meter set close to the sample. The intensity values were taken from the spectrometer, based on 20 accumulations of 5s exposures. The broad Gaussian nonlinear intensity curves are the result of the 33\text{nm} laser linewidth, typical of a tunable Ti:Sapphire system with a broad emission spectrum. Figures 4.15.(c,f) were generated by taking the peak SHG and THG intensities from the spectrometer.
data, and plotted against the incident power using a log-log scale. From the least square fit on the log-log scale plot, we have obtained the SHG and THG slopes 1.98 and 2.87 for the transmissive sample, and 2.00 and 2.60 for the reflective sample. Within the experimental error, the results suggest that the SHG and THG processes have a quadratic and cubic dependence with the incident power.
Figure 4.15: SHG and THG measurements as function of the incident power and at normal incidence $\theta=0^\circ$. (a) corresponds to a measurement of the transmissive sample using incident polarization $\psi = 0^\circ$ and excitation wavelength $\lambda_1^T = 1888\,nm$, and (b) corresponds to the reflective sample with $\psi = 90^\circ$ and $\lambda_2^R = 1816\,nm$. (c,f) Power dependence measurements of the (green) SHG and (blue) THG signals from the transmissive and reflective sample respectively. Data points were obtained from the peak intensity values in plots (a,b,d,e) and power meter data.
Chapter 4. Tunable bulk nonlinear effects using complementary metamaterials

<table>
<thead>
<tr>
<th>Sample</th>
<th>SHG (cts/s)</th>
<th>THG (cts/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>transmissive</td>
<td>62</td>
<td>162</td>
</tr>
<tr>
<td>reflective</td>
<td>19</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 4.3: Comparison of SHG and THG intensities at maximum average power (100mW) between complementary samples.

The peak intensities at the maximum average power (100mW), obtained from the fitting, are presented in Tab. 4.3, and show that the transmissive sample provides a larger enhancement of the nonlinear processes– $3 \times$ for SHG and $16 \times$ for THG.

### 4.6.3 Reference sample measurements

![SHG and THG measurement](image)

Figure 4.16: SHG and THG measurement on two reference samples of (a) transmissive and (b) reflective nanostructures using three different incident polarizations $\psi = [0^\circ, 45^\circ, 90^\circ]$. Both samples were excited at normal incidence $\theta = 0^\circ$ with a 100 mW average power and fixed wavelength $\lambda_1^T = 1879$ nm for the transmissive sample and $\lambda_1^R = 1884$ nm for the reflective sample. Data collected over 9 accumulations of 5 s exposures.

Following the process described in Fig. 4.3, we have fabricated two complementary samples of asymmetric cross arrays without the ZnO film (the last MBE
The LSP resonances of these samples are $\lambda_1=1828\text{ nm}$ and $\lambda_2=2709\text{ nm}$, which are slightly higher than expected due to variations in the nanostructure dimensions. We have performed SHG and THG measurements on these samples using the setup described in Fig. 4.14. The results are presented in Fig. 4.16 where (a) corresponds to a sample of transmissive crosses and (b) reflective crosses—details on the experimental conditions are presented in the image caption.

No SHG contributions were observed in the measurement, suggesting that surface SHG from the metallic structures are below the noise level of the CCD detector. Some THG effects can be seen in particular in the reflective sample, and the change in incident polarization seems to suggest control over this effect. The highest THG intensity measured in both samples was 0.7 cts/s which, in comparison to the power measurements values in Tab. 4.3, is very close to the noise level of the detector.

Figure 4.17: Nonlinear contributions from (black) bare $\text{Al}_2\text{O}_3$ substrate and (red) $\text{ZnO}$ film deposited on $\text{Al}_2\text{O}_3$—no nanostructures on either of the samples. Measurement was achieved at normal incidence $\theta=0^\circ$ with a 75 mW average power, 1868 nm incident wavelength and fixed arbitrary polarization $\psi$. Data collected over 12 accumulations of 5 s exposures.
From the reference samples it is unclear what material contributes to THG. An additional measurement was performed on both Al₂O₃ and ZnO deposited on Al₂O₃, without the nanostructures. The result from this experiment is presented in Fig. 4.17 and it shows no signs of THG or SHG from these samples, strongly suggesting that bulk nonlinear effects on the ZnO can only be activated with the plasmonic nanostructures.

### 4.6.4 Polarization dependence

We now analyse the nonlinear optical response from the complementary nanostructures with the ZnO film, as function of the incident electric-field polarization $\psi$, at normal-incidence $\theta=0^\circ$ and with the pump wavelength fixed to the lowest plasmonic resonance—$\lambda_1^T=1872$ nm for the transmissive sample and $\lambda_1^R=1803$ nm for the reflective sample. Figures 4.18.(a,b) show the background corrected and normalized SHG and THG intensities from the transmissive sample, and likewise, Fig. 4.18.(d,e) shows the same effects for the reflective sample. The results were obtained from a 30 s exposure to a 100 mW beam (average power) at an 80 MHz repetition rate. The gaussian profile of the normalized intensities are due to the $\sim 22$ nm laser linewidth at the fundamental wavelength. From the gaussian curves, we have then extrapolated the peak SHG and THG intensities and plotted this as function of the incident electric-field polarization $\psi$—this result is shown in Fig. 4.18.(c) for the transmissive sample and Fig. 4.18.(f) for the reflective sample. In these two plots, the linear transmission data from the FTIR measurement in Fig. 4.5.(a,d) was also included.

As discussed above, the rotation of the incident electric-field polarization $\psi$ has an effect on the intensity of the two plasmonic modes inside the cross. In the linear regime, this effect can be described by the function $A \cos^2 \psi + B \sin^2 \psi$, where $A$ and $B$ are the transmission values for polarizations $\psi=0^\circ$ and $90^\circ$ respectively, and are essentially a measure of the intensity of the horizontal and vertical plasmonic modes at fundamental excitation. The fitting of the nonlinear results is however not as simple, since the nonzero tensor components of the nonlinear susceptibility

116
Figure 4.18: (a,d) SHG and (b,e) THG measurements as function of incident electric-field polarization $\psi$ for the (a,b,c) transmissive and (d,e,f) reflective samples. At normal incidence $\theta = 0^\circ$ and with linearly polarized light, devices were excited at plasmonic resonance $\lambda_1$ and with a 100mW average power. Plots (c) and (f) were extrapolated from the peak values obtained from plots (a,b) and (d,e), and FTIR data in Fig. 4.5.(a) and 4.5.(d). Solid lines correspond to the power $S(\omega_g)$ obtained from Eq. 4.3 using a multipole expansion of the linear, SHG and THG polarizations (Eq. 4.1a, 4.1b and 4.1c) for excitation of the plasmonic mode $\lambda_1$ (as defined in Fig. 4.6) and varying incident electric-field polarization $\psi$. Dashed red line corresponds to a fit of the function $A \sin^2 \psi + B \cos^2 \psi$ with $A=1$ and $B=0.19$ for transmissive sample, and $A=1$ and $B=0.13$ for the reflective sample.
of the ZnO film plays a role in defining the contribution of the different plasmonic modes, i.e. the nonlinear response of the ZnO film is sensitive to the amplitude of the different electric field components \((E_x, E_y, E_z)\), as described in Eq. 4.1b and Eq. 4.1c. Therefore, as we change the incident electric-field polarization and couple to different LSP modes in the cross, different nonlinear tensor components are excited leading to a complex nonlinear optical response.

Using the multipole expansion presented in Sec. 4.4, the total power and individual multipole contributions were calculated for the linear and nonlinear processes as function of the incident polarization \(\psi\). In these calculations, we have taken the lowest LSP mode, \(\lambda^T_{LSP} = 1866\,\text{nm}\) for the transmissive sample and \(\lambda^R_{LSP} = 1828\,\text{nm}\) for the reflective sample (Fig. 4.6). The numerical results for the total power are given in Fig. 4.18.(c,f) by the solid lines. We see a clear agreement between experimental and numerical data for the linear calculations (red line), with discrepancies coming from slight differences in geometry and material properties between model and fabricated devices. With regards to the nonlinear results, as we rotate the incident polarization towards the angle with strongest plasmonic enhancement (\(\psi=0^\circ\) for the transmissive sample and \(\psi=90^\circ\) for the reflective sample), both the experimental and numerical data show an increase in nonlinear activity. This behaviour is complementary between the transmissive and reflective samples. However, the evolution of the curves as we reach a stronger enhancement is not identical.

We now move to explain the differences in the total power between complementary samples, using the numerical method introduced in Sec. 4.4 based on a multipole expansion of the linear and nonlinear polarizations. Starting with SHG, in Fig. 4.19.(a) and Fig. 4.19.(c) we plot the total power and individual multipole contributions for varying electric-field polarizations \(\psi\), for the transmissive and reflective samples respectively. The electric dipole is the main process for SHG in both samples, however, a small contribution from the magnetic dipole and electric quadrupoles is visible when the phase difference allows a certain level of constructive interference. The phase difference is independent on polarization for the transmissive sample, while for the reflective sample, Fig. 4.19.(c) suggests
Figure 4.19: Total lattice power and separate multipole contributions as function of incident polarization $\psi$, from the (left) transmissive and (right) reflective samples for (top) SHG and (bottom) THG nonlinear processes. In the model, the source was set to normal incidence $\theta=0^\circ$, and wavelengths $\lambda_T^1=1866 \text{ nm}$ and $\lambda_R^1=1828 \text{ nm}$ for the transmissive and reflective samples (respectively). Data normalized to the maximum total power of each individual nonlinear process and device.

stronger constructive interference between magnetic dipole and electric quadrupole, at polarizations closer to resonance ($\psi=90^\circ$). The difference in SHG response between complementary samples can be explained by realising that the magnetic dipole and electric quadrupole have a dependence on the spatial profile of the electric-field, given by the product between the position $\mathbf{r}$ and polarization $\mathbf{P}(2\omega, \mathbf{r})$ vectors (Eq. 4.2b and 4.2c). It is therefore expected that variations in the electric-field distribution between complementary samples (Fig. 4.9 and
4.10), will have a significant impact on the evolution of the SHG power as we increase the field strength inside the nonlinear medium. With regards to the reflective sample data in Fig. 4.9.(c), the intensity of the electric-quadrupole contribution is twice the value of the combined multipole power but due to a phase difference with the magnetic dipole, this does not translate in a stronger SHG signal.

With regards to the THG data, our numerical calculations in Fig. 4.19 suggest that the shape of the total power curve is mainly influenced by electric-dipole contributions. Magnetic-dipole and electric-quadrupole contributions are actually much higher in amplitude by 6 orders of magnitude (insets). However, due to destructive interference they do not play a significant role in the THG process. Equal magnitude of these contributions has also been reported in Basharin [176]. The same cannot be assumed for the experimental result, since the quality of the ZnO films are far from the ideal case assumed in the simulations. There is a possibility that higher order-multipoles (e.g. octupoles) can play a role in the THG since this is a smaller wavelength. However, we have truncated the number of terms up to second order multipoles for simplification purposes. Focusing on the electric-dipole result, to explain the different evolution of the THG power between complementary samples, we need to look back at the definition of the THG polarization in the ZnO (Eq. 4.1c). Essentially, the x and y components of \( P(3\omega) \) have a dependence on both \( E_x(\omega, r) \) and \( E_y(\omega, r) \) electric-fields, and therefore, the two LSP modes at each arm of the cross can contribute simultaneously to the THG process, with different intensities controlled by the incident polarization \( \psi \). Differences in the curve evolution can then be explained by the higher plasmonic losses in the transmissive sample, leading to weaker off-resonance LSP modes. This effect is further amplified by the quadratic and cubic dependence of the THG polarization to the off-resonant electric-field components, namely \( E_y(\omega, r) \) for the transmissive sample and \( E_x(\omega, r) \) for the reflective sample (Eq. 4.1c). Another observation from this result is the local minimum between complementary structures, \( \psi=30^\circ \) for the transmissive sample and \( \psi=60^\circ \) for the reflective sample, which corresponds to the angle of weakest simultaneous excitation of the LSP modes in each arm of the cross.
Coming back to the analysis of the experimental data, we see some discrepancies of the nonlinear processes with the numerical results. First, a few assumptions were made in the model regarding the nonlinear susceptibility of the ZnO films (Sec. 4.4). The nonlinear susceptibility of a material is highly dependent on the arrangement of atoms in its crystal structure. Parameters such as thickness of the film, grain size, crystallinity, surface quality, and deposition method, are known for playing a large role in the absolute value of $\chi^{(2)}$ and $\chi^{(3)}$. Therefore, the values used in the model might not be an accurate representation of the actual films [111]. Furthermore, both the SHG and THG polarizations have contributions that are dependent on the rotation of the crystal c-axis, namely $\chi^{(2)}_{15} = \chi^{(2)}_{24}$ and $\chi^{(3)}_{18} = \chi^{(3)}_{29}$. In the model, the xy plane of the epitaxial ZnO film was assumed to be aligned with the plasmonic nanostructures. In reality, the ZnO film takes the orientation with minimum interface energy, which typically is dependent on the substrate interface and there is no selectivity during the growth. This in turn has an impact on the crystallinity of the ZnO film, and given the different interfaces, it is expected to affect the nonlinear susceptibility values between complementary samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\psi$(deg)</th>
<th>SHG (cts/s)</th>
<th>THG (cts/s)</th>
<th>SHG (dB)</th>
<th>THG (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>transmissive</td>
<td>0</td>
<td>24</td>
<td>55</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reflective</td>
<td>0</td>
<td>1.1</td>
<td>0.7</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>20</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Comparison of SHG and THG intensities between complementary samples for two incident polarizations $\psi=[0^\circ,90^\circ]$ and an average power of 100 mW. Attenuation values calculated using the expression $Att$ (dB) = $10 \times \log_{10}(\max[I(\psi)]/\min[I(\psi)])$. 

121
Table 4.5: Modelled transmitted SHG and THG power for two incident electric-field polarizations $\psi=[0^\circ,90^\circ]$ through (row 1) a $\text{Al}_2\text{O}_3$ substrate with an unstructured 75nm ZnO film, (row 2) through the transmissive sample and (row 3) through the reflective sample. The source was set to normal incidence $\theta=0^\circ$, electric-field amplitude $|E|=1 \text{ V/\mu m}$, and wavelength $\lambda=1904\text{ nm}$. Rows 4 and 5 show the enhancement owning to the LSP modes for the transmissive and reflective samples respectively. This is in relation to the measurements in row 1 i.e. ratio between sample and unstructured ZnO film. Rows 6 and 7 show the enhancement due to the change of the incident polarization $\psi$. Rows 8 and 9 show a comparison of the nonlinear response between complementary samples for polarizations with highest and lowest enhancement respectively.

The measured intensities from the CCD detector are presented in Tab. 4.4. For the same laser power (100 mW), the attenuation of the nonlinear signals is similar in value between complementary samples, with the exception of SHG in the transmissive samples which is slightly lower, and averaging to a 11.25 dB attenuation. As already observed in the power measurements (Sec. 4.6.2), the larger field enhancement is obtained with the transmissive sample. Differences in the tabled results for the transmissive device (Tab. 4.3 and Tab. 4.4), are due to a slight misalignment of the sample with regards to the focal point of the laser beam.
A more detailed comparison in performance can be obtained from the modelled results presented in Tab. 4.5. Here, the case of a 75nm unstructured ZnO film has also been considered. Without the metallic nanostructures (bare ZnO), the nonlinear processes are the sole result of an electric-dipole contribution, i.e. \( S_{\text{md}}(2\omega) = S_{\text{eq}}(2\omega) = 0 \). Taking into account Eq. 4.1b, it can be easily found that an unstructured ZnO film cannot support SHG at normal incidence, since \( E_z = 0 \) and \( P_z(2\omega) \) is not seen by the detector (\( n = [0, 0, 1] \) ⇒ \( S_{\text{ed}} \propto [P_x, P_y, 0] \) in Eq. 4.3). With the introduction of the metasurface, the electric-field vector gains a nonzero z-component (Fig. 4.10) allowing the excitation of the the x and y components of the SHG polarization. THG can however be excited at normal incidence on bare ZnO since \( P_x(3\omega), P_y(3\omega) \neq 0 \) (Eq. 4.1c). Based on the tabled results, an enhancement of \( 102.6 \times \) and \( 17.0 \times \) can be achieved with the transmissive and reflective samples respectively. We explain the reason for no observation of THG from the unstructured ZnO film sample (Fig. 4.17) as result of insufficient detection capability, and small volume of film not allowing a sufficient build-up of signal.

Also in Tab. 4.5 and rows 6 to 7, we also compare the difference in nonlinear response due to the change of incident polarization. For SHG the reflective sample offers the best attenuation of the two samples, in agreement with experimental data, while for THG the transmissive sample offers the best attenuation. We also make a comparison between complementary samples, as shown in rows 8 to 9, and our results suggest a stronger SHG signal from the reflective sample, and a larger THG from the transmissive sample.

## 4.6.5 Wavelength dependence

We have also looked at the wavelength dependence of the nonlinear processes for three different incident polarizations \( \psi = [0^\circ, 45^\circ, 90^\circ] \), normal incidence \( \theta = 0^\circ \), and a fixed average source power of 60 mW. Due to restrictions in the lasing power for wavelengths above 2000 nm (Fig. 3.14), and the significant decrease in detection
efficiency for \( \lambda > 1000 \) nm, we have limited ourselves to study frequencies around the first plasmonic mode \( \lambda_{1}^{T,R} \), i.e. within the range \( 1750 < \lambda < 1950 \) nm. In the next two sections we will present the SHG and THG results from this experiment, and include for the SHG case a more detailed numerical analysis where the wavelength dependence of the second plasmonic mode \( \lambda_{2} \) is also observed.

**Second harmonic generation**

Figure 4.20 shows both numerical and experimental results for the (top) linear and (bottom) SHG light transmitted across the (left) transmissive and (right) reflective samples. The solid and dashed lines in the linear plots, and the dots and dashed lines in the SHG plots, represent experimental and numerical results respectively. We have only introduced the linear results for comparison with the SHG, and they are essentially a zoom-in (around the \( \lambda_{1} \) resonance) of the FTIR and numerical data in Fig. 4.5 and Fig. 4.6.

The SHG result in Fig. 4.20.(b) for the transmissive sample and Fig. 4.20.(d) for the reflective sample, clearly shows a strong dependence of the SHG process to both the wavelength of excitation \( \lambda \) and the incident polarization \( \psi \). The experimental data given by the dots has been corrected to take into account the wavelength dispersion of the grating in the spectrometer and the CCD detector. There is a reasonable agreement between experimental and numerical data (dashed lines), given by the evolution of the SHG with the pump wavelength \( \lambda \), relative amplitude between polarizations \( \psi \), and the blue-shift in the SHG peak with respect to the harmonic of the LSP resonance. In our numerical model, the blue-shift was to be expected since \( S(\omega_{g}) \propto \omega_{g}^{2} \) (Eq. 4.3), and therefore, the SHG peak is shifted towards shorter wavelengths, in this case by 114 nm for the transmissive sample and 76 nm for the reflective sample. Discrepancies between the experimental and numerical data for the SHG process can be explained by the assumptions made regarding the nonlinear susceptibility \( \chi^{(2)} \) of ZnO, and also due to inaccuracies in detecting photon energies closer to 1000 nm wavelengths, mainly at off-resonance angles \( \psi \) where weaker SHG are produced.
Figure 4.20: Spectral measurement of the (a,c) linear (solid lines) and (b,d) SHG intensity (dots) for three different incident polarizations $\psi = [0^\circ, 45^\circ, 90^\circ]$, normal incidence $\theta = 0^\circ$ and pulse power 60mW. Left column corresponds to data from the transmissive sample, and right the reflective sample. Dashed lines are the results from the numerical calculations.

A more complete analysis of the SHG can be seen in Fig. 4.21.(b,e) and Fig. 4.22, corresponding to the total lattice power and individual multipole contributions respectively. In these numerical calculations, we have included the effect of both LSP resonances on the SHG response. Starting with the sum of multipole contributions $S(2\omega)$ from Eq. 4.3, we see that both samples are characterized for two SHG peaks ($\lambda_{1T,R}^2/2$ and $\lambda_{2T,R}^2/2$) which are dependent on polarization and wavelength. At resonant wavelength, the polarization state between complementary samples for the SHG enhancement is phased by $\Delta\psi = 90^\circ$, as predicted by Babinet’s
principle. Also, we observe a blue-shift of the two SHG peaks due to $S(\omega_g) \propto \omega_g^2$ in our formulation of the lattice power. There are, however, some significant differences in the SHG response between complementary samples, the most obvious being the loss of the isosbestic point– the wavelength at which transmission is invariant with polarization– in the reflective sample (Fig. 4.21.(e)). Also, the $\lambda_2^{T,R}/2$ peak is much broader in the reflective sample which we believe is due to the truncated multipole expansion. Essentially, the $\lambda_2$ resonance originates from an LSP mode in the horizontal arm of the complementary crosses (Fig. 4.9), corresponding to the smallest feature of this geometry ($W_H=70\text{nm}$). This is two orders of magnitude smaller than the SHG wavelength, and therefore might be susceptible to higher-order multipole effects.

A closer look into the multipole decomposition of the SHG in Fig. 4.22, reveals that the complementary samples are characterized for different contributions. We have normalized the results with respect to the maximum SHG power for comparison purpose. The electric-dipole $S_{ed}(2\omega)$ is the most dominant process in both samples. However the reflective sample has stronger magnetic-dipole $S_{md}(2\omega)$ and electric-quadrupole $S_{eq}(2\omega)$ interactions in the order of $1.5 \times$ and $1.8 \times$ the total SHG power $S(2\omega)$ respectively, which, due to destructive interference, play a small role in the SHG. In comparison to the linear results in Fig. 4.6, the $S_{md}$ and $S_{eq}$ are 1 to 2 orders of magnitude higher in the SHG process.
Figure 4.21: Complete numerical spectral analysis of (a,d) linear and (b,c,e,f) SHG effects from the (left) transmissive and (right) reflective samples for a range of incident polarizations $\psi$. Bottom row corresponds to the SHG power after accounting for the linear optical response of the metastructures at SHG wavelengths, as shown in Fig. ??.
Figure 4.22: Multipole decomposition of the SHG process from the (left) transmissive and (right) reflective samples, and for a range of incident polarizations $\psi$. Top, middle and bottom plots correspond to SHG power contributions from the electric-dipole $S_{ed}(2\omega)$, magnetic-dipole $S_{md}(2\omega)$ and electric-quadrupole $S_{eq}(2\omega)$ respectively, normalized to the total radiated power $S(2\omega)$ in Fig. 4.21.(b) and Fig. 4.21.(e).
Another important aspect in the description of the nonlinear processes is linear response of the metasurfaces at nonlinear wavelengths [181]. This process can be accounted by calculating the linear power at those wavelengths using Eq. 4.1a and Eq. 4.3, which results have been presented in Fig. 4.8. The global response can then be derived by multiplication of the linear power at SHG wavelengths, and the SHG power from the nonlinear process at fundamental wavelengths. This result is presented in Fig. 4.20.(c) and Fig. 4.20.(e) for the transmissive and reflective samples respectively, including two incident electric field polarizations $\psi = [0^\circ, 90^\circ]$. These results show a large attenuation of the SHG at wavelengths around the second resonance $\lambda_2$, owning to the poor linear power at these wavelengths. The features highlighted in Fig. 4.20.(c) for $\psi = 90^\circ$ are originated from the linear response in Fig. 4.8.(a).

**Third harmonic generation**

We have also attempted a wavelength dependence measurement of THG as shown by the dots in Fig. 4.23.(b) and Fig. 4.23.(e) for the transmissive and reflective samples respectively, using three different polarizations $\psi = [0^\circ, 45^\circ, 90^\circ]$. In these results, we have taken into account the wavelength dispersion of various optical components in our setup, in particular the monochromator grating and CCD detector as shown in Fig. 3.15.(b) and Fig. 3.16. Although the THG measurement suggests a wavelength and polarization dependence, it is difficult to analyse the results objectively. The laser and detector are being operated in a frequency range where they are most unstable, and since the THG has a cubic dependence to the incident field, this can lead to large experimental noise. Also to note that THG can also be generated in other parts of the system, since it is not dependent on the symmetry of a material. A full numerical analysis of the total transmitted THG power, including enhancements from the two LSP modes, is shown in Fig. 4.24. Similarly to the polarization results in Sec. 4.6.4, we found that the wavelength dependence of the THG process is also characterized for electric-dipole contributions.

The effect of the linear response at THG wavelengths has been accounted in the
Figure 4.23: Spectral measurement of the (a,c) linear (solid lines) and (b,d) THG intensity (dots) for three different incident polarizations $\psi = [0^\circ, 45^\circ, 90^\circ]$, normal incidence $\theta = 0^\circ$ and pulse power 60mW. Left column corresponds to data from the transmissive sample, and right the reflective sample. Dashed lines are the results from the numerical calculations.

The global response from the metasurfaces, similarly to what has been done in the SHG case. These results are presented in Fig. 4.24.(c) and Fig. 4.24.(f) for the transmissive and reflective samples respectively. Across the THG range, the transmitted linear power decreases with increased wavelength (Fig. 4.8). However, it does not reach the lower end as observed for larger SHG wavelengths near the second LSP resonance $\lambda_2$. As result, the model suggests that both LSP modes can have contributions in the THG response. For the transmissive sample, however, three features arise around the second LSP resonance originated from the linear response at $\psi = 90^\circ$. 
Figure 4.24: Complete numerical spectral analysis of (a,d) linear and (b,c,e,f) THG effects from the (left) transmissive and (right) reflective samples for a range of incident polarizations $\psi$. Bottom row corresponds to the THG power after accounting for the linear optical response of the metastructures at THG wavelengths, as shown in Fig. 4.8.
4.7 Conclusion

In summary, we have presented a detailed experimental and numerical study of bulk SHG and THG nonlinear effects from a ZnO film with enhancements from complementary asymmetric cross metasurfaces. ZnO has both second- and third-order nonlinear susceptibilities which, combined with a large transparency range, good thermal properties and range of deposition techniques, makes it a versatile material for integrated photonic devices. A simple fabrication process using EBL, argon-milling and MBE (Fig. 4.3) was used to produce complementary metasurfaces covered with a ZnO film on an insulating substrate. The method can be easily adapted to accommodate different plasmonic structures, substrate materials, or deposition techniques. MBE growth has the advantage of producing high-quality crystalline films (Fig. 4.4), but care needs to be taken in the choice of substrate, as this will dictate the orientation of the film based on the lattice mismatch of lowest energy. We found very different film growths between the continuous and discontinuous nanostructured gold films (transmissive and reflective samples). Both samples were grown with c-axis oriented (0002) ZnO films. However, the transmissive sample shows the better crystallinity of the two films (Fig. 4.4), and therefore we expect the nonlinear susceptibility values of the ZnO to differ between samples.

The linear optical response obtained from the FTIR measurements (Fig. 4.5) and multipole calculations (Fig. 4.6), in conjunction with the EM-field profile studies at fundamental wavelength (Fig. 4.9 and Fig. 4.10), further validated the applicability of Babinet's principle for the case of asymmetric cross-shaped nanostructures. These studies also helped elucidate the role of the LSP resonances on the field enhancements around the metamaterials, and the key electric-field contributions to the nonlinear processes in the ZnO.

Our nonlinear experimental and numerical results successfully show bulk SHG and THG processes from ZnO films, which previously has been difficult to attain due to the small birefringence of this material. Measurements with reference
samples consisting of a ZnO film grown on Al₂O₃ without the nanostructures, and the transmissive and reflective samples without the ZnO film, show no nonlinear activity from the samples within the detection range. The electric-field enhancements from LSP excitations are therefore the key effect that allows to leverage the nonlinear properties of thin ZnO films without the need of phase-matching. We have shown tunability of the nonlinear response by controlling the incident electric-field polarization and fundamental wavelength. A complex nonlinear optical response was observed owing to the interaction between different LSP modes in the asymmetric cross (associated with \( E_x \), \( E_y \) and \( E_z \)), and tensor components of the nonlinear susceptibilities. Between complementary nanostructures, larger nonlinear effects were observed from the sample with transmissive crosses with enhancements on the order of \( 2.2 \times \) for SHG and \( 10.6 \times \) for THG, based on averages calculated from Tab. 4.3 and Tab. 4.4. Similar comparison was made using the numerical model of the metasurfaces with results within the same order of magnitude (Tab. 4.5). The multipole expansion of the nonlinear polarizations reveals a strong contribution from the electric-dipole for both SHG and THG, although magnetic-dipole and electric-quadrupole have a larger relative value. Due to destructive interference at the far-field, the higher-order multipoles are largely unexplored in cross-shaped nanostructures, opening up the possibility to investigate alternative geometries based on metal-dielectric-metal interfaces [131]. It was also found that the linear response of the metasurfaces at SHG and THG wavelengths have a large influence in the global nonlinear response. In particular for the SHG case, it was found a significant attenuation of the SHG signal around the second LSP resonance. Further optimization of the metastructures linear response at nonlinear wavelengths is then required to minimize significant attenuation of the nonlinear signals. This work will find applications in integrated photonic systems, requiring an optical-control mechanism for wavelength conversion and filtering, or lasers and OPOs where there is a need to optimize the nonlinear processes.
Chapter 5

SHG-CD response from superchiral metasurfaces

5.1 Introduction

Chirality is a structural property given to molecules when their mirror image (enantiomer) cannot be superimposed onto itself. Circularly polarized light (CPL) - also of chiral nature - will experience a different complex refractive index ($\tilde{n} = n + i\kappa$) when travelling across a chiral molecule, resulting in (1) a change of speed (for $n \neq 0$) and/or (2) different degrees of absorption and light scattering (for $\kappa \neq 0$) depending on the handedness of the incident light. These effects are called optical rotation and circular dichroism (CD) respectively and have seen applications in molecular spectroscopy [182]. Linear chiral optical effects have their origin in the coupling between electric dipole transitions and either magnetic dipole or electric quadrupole transitions [13].

Chiral optical effects in chiral molecules are of a small magnitude due to the molecular bond length being of a much smaller scale (hundreds of picometers) than the helical pitch of visible CPL (hundreds of nanometers). This effect can be quantified using the dissymmetry factor $g^*$ by looking at the difference of light absorption between left- and right CPL excitations. Typical values in most molecules are in the order of $|g| \sim 10^{-3}$ for visible wavelengths [183]. An increase
of the dissymmetry factor can be achieved by (1) increasing the chiral pitch of the material to a size comparable to the wavelength of CPL light (impractical and limited to a single compound), or by (2) compressing the helical pitch of CPL to match the molecular bond length of the material. With recent advancements in plasmonic research and nano-fabrication, it is now possible to realize and fabricate metallic nanostructures with an EM response with a twist far more pronounced than of CPL light (superchiral light). This is extremely useful in molecular spectroscopy and sensing applications as it offers a more accurate and sensitive measurement, allowing the use of smaller sample quantities [184].

As we have seen in Sec. 2.1, interesting optical effects can be observed at the surface of metals when light couples with the free electron oscillations in a metal-dielectric interface due to the existence of SPs. In the case of metallic nanoparticles, strong EM resonances can be observed due to an increase of the surface-to-volume ratio of the particle and confinement of the free electrons. At resonance, the surface currents in the nanoparticles will vary with the EM field of the incident light, and their propagation at the surface is dependent on the particle geometry. Enantioselectivity and enhancement of chiral optical effects can be achieved with plasmonic nanoparticles with certain geometries or by changing the metal-dielectric interface with the introduction of chiral molecules. Many applications can be found in chiral plasmonic nanoparticles, such as in ultrasensitive detection of chiral biomolecules [184], negative refraction [185], broadband circular polarizers [186] and nanolevitation by repulsive Casimir forces [187].

In Sec. 2.2, we have seen how intense optical fields, comparable to the fields that bind valence electrons to the nucleus of atoms, can induce a nonlinear polarization in a material changing their linear optical properties. We have also looked into the particular case of SHG processes and we have seen that this can only occur in noncentrosymmetric materials or at the surface of materials where the centrosymmetry is broken. Chiral molecules, due to their atomic arrangement, have also no mirror or inversion symmetry which equates to additional tensor components in the representation of the second-order susceptibility $\chi_{ijk}^{(2)}$ (chiral
and achiral/surface components). In contrast to the linear effects, nonlinear chiral optical effects are usually associated with electric dipole transitions. It has been shown that magnetic-dipole and electric-quadrupole contributions can enhance surface SHG in chiral molecules [188]. Also, chiral measurements from nonlinear processes have a larger contrast than their linear counterpart, and so, these are extremely useful in spectroscopy techniques for studying chemical and biological systems [13].

In this chapter, we investigate a type of plasmonic nanostructure with enhanced chiral optical properties. A plasmonic device has been realised with a tunable superchiral optical response. Mapping of the superchiral fields with areas of strong SHG activity has been achieved using a combination of experimental and numerical methods. This chapter is divided in three parts: Sec. 5.2 contains a description of the device, including the reason for interest in gammadion structures in the study of chiral optical effects; Sec. 5.3 briefly discusses the main SHG results obtained using a confocal microscope; and in Sec. 5.4, details on the numerical work used for mapping regions of superchiral light and SHG polarization are presented.

5.2 Gammadion L-shaped devices with tunable SHG-CD response

The device is composed of 4 L-shaped nanostructures which in a periodic arrangement form 2 gammadion cross unit cells of different chiralities- clockwise $\mathcal{H}$ and anti-clockwise $\mathcal{F}$ - as shown in Fig. 5.1. The width and length of the L-shapes are fixed to 200 nm and 1 $\mu$m respectively. The nanostructures are made of 30 nm of gold with a 5 nm chromium adhesion layer, and are sitting on 100 nm of silicon dioxide (SiO$_2$) grown via thermal annealing on an Si substrate. CPL light hits the nanostructured side of the sample at normal incidence, and the reflected light is collected for measurement.

The devices are patterned using EBL by spin coating PMMA positive electron
beam resist. After development of the resist in an MIBK:IPA solution, the chromium and gold layers are thermal evaporated onto the sample, following a lift-off step in acetone which removes the metal layers and resist around the nanostructures. The EBL design is set so the central distance of the L-shapes forming the $\mathcal{F}$ unit cell is fixed at a distance around 200 nm. For the $\mathcal{F}$ unit cell, various devices were produced with different central distances $d$. Top view SEM micrographs of the $\mathcal{F} / \mathcal{P}$ unit cells are shown in Fig. 5.2.

Figure 5.1: Schematic depiction of the L-shaped, gammadion arranged, periodic nanostructures forming $\mathcal{F} / \mathcal{P}$ unit cells. Main dimensions and material thickness’s are also depicted in the 3D model.

Figure 5.2: Composite SEM micrograph of ten different L-shaped, gammadion arranged, periodic nanostructures. The arrows indicate the varying $d$ distance between $\mathcal{F}$ unit cells, and fixed distance for $\mathcal{P}$ unit cell. Figure reproduced from Valev et al. [70].

Strong linear chiral optical activity has already been reported for gammadion type plasmonic nanostructures [189]. Here we look at the nonlinear chiral optical effects using a design that contains a dissymmetry between $\mathcal{F} / \mathcal{P}$ unit cells, with the objective of enlightening the role of the local enhancements of superchiral light in the different multipole components of the SHG polarization.
5.3 Experimental methods and results

For the near-field measurement of SHG from the devices in Fig. 5.2, we have used a confocal laser scanning microscope (Zeiss LSM 510 META) whose operating principles have been discussed in Sec. 3.2.6. The devices were excited using a femtosecond pulsed Ti:Sapphire laser at a 900 nm fundamental wavelength, close to an LSP resonance in the nanostructures (Sec. 5.4.1). Light was then focused on the nanostructure side of the sample at normal incidence, and with a spot size of 580 nm (FWHM at 1/e^2) using a Zeiss 40x Plan-APOCHROMAT objective. Light reflected from the sample was then directed to a dichroic mirror (HFT KP650) and a short-pass filter (KP685) which removed any fundamental signal. SHG light was finally collected by an PMT, and images formed by scanning the device array with a pixel dwell time of 102.4 µs.

Figure 5.3 shows the confocal microscope measurements, where the brighter spots represent regions of high SHG intensity. In these images, the nanostructures have the same arrangement as in Fig. 5.2 where the top-row represents excitation with right-CPL light (or ⟳) and the bottom row with left-CPL light (or ⟲). As highlighted by the red and yellow dashed squares, the formation of the SHG spots coincides with the centre of the gammadion with the same chirality as the CPL. The intensity, however, varies across nanostructures with differing central distance d.
Figure 5.3: SHG microscopy measurement of the devices in Fig. 5.2. Incident laser excitation at 900 nm wavelength with right-CPL (top) and left-CPL (bottom). Red and yellow dashed squares highlight regions of strong SHG intensity corresponding to the centre of the gammadion structure with the same chirality as the incident CPL. Figure reproduced from Valev et al. [70].
A more quantitative analysis of the results obtained above is presented in Fig. 5.3. Figure 5.3.(a) shows the SHG intensity ratio between left- and right-CPL excitation on the gammadion with \( \Sigma \) and \( \Pi \) chiral centres at different gap distances \( d \). The intensity values were obtained by integrating the pixel intensity inside the dashed squares. The result shows a nearly constant SHG-CD from the nanostructure with chiral centre \( \Pi \) while its enantiomer \( \Sigma \) has a non-monotonic variation with the gap size. Further evidence of the SHG-CD is given by the separation distance of the yellow and red lines from the white line representing equal intensity at different light handednesses, i.e. \( I^L(2\omega)/I^R(2\omega) = 1 \). This result suggests that there is a non-local bulk effect that contributes to the SHG process, adding to the surface SHG polarization which relies on local field enhancements at the metal surface only.

Figure 5.4: (a) SHG dichroism from \( \Sigma \) unit cell (yellow) and \( \Pi \) unit cell (red), calculated through the ratio between the integrated intensity with left- and right-CPL for different gap distances \( d \). (b) ratio of SHG intensity \( I^L(2\omega)/I^R(2\omega) \) as a function of gap size \( d \). The lines are guides to the eye. Figure reproduced from Valev et al. [70].

Figure 5.4.(b) shows the intensity ratio \( I^L(2\omega)/I^R(2\omega) \) as function of the centre distance \( d \) for the \( \Sigma \) unit cell, verifying that for the case when \( d \approx 207 \text{ nm} \), the SHG intensity is equal in both unit cells since their gap distance is nearly identical. From this plot we can see an SHG-CD peak when \( d = 253 \text{ nm} \) which is close to half the SHG wavelength \( \lambda_{2\omega}/2 = 225 \text{ nm} \). This peak could be due to a constructive
interference of the SHG light emitted by opposite interfaces.

To summarize, these results show tunability of the SHG-CD by changing the gap distance of the unit cell with centre. Variation of the integrated SHG intensity we believe is due to an added (non-local) component in the description of the SHG polarization, namely electric quadrupole contributions. In the next section, a more rigorous investigation of the linear and nonlinear response in these nanostructures will be carried out by means of simulation tools and modelling techniques, in particular the mapping of near-field enhancements with regions of superchiral light and strong SHG response.

5.4 Numerical calculations

For the numerical calculations presented in this chapter, we have used a commercial software, Synopsys® DiffractMOD™ version 2013.12-2, to obtain the EM field distribution at the fundamental frequency of the nanostructures presented in Sec. 5.2. This software employs the RCWA method discussed in Sec. 3.3.1 to solve EM scattering problems in periodic structures. Dispersion of the dielectric constants in the metals was taken into account by means of the Lorentz-Drude model [31]. In this section, we will start with a linear optical analysis of the L-shaped gammadion nanostructures- i.e. spectral response, EM field distribution, and linear polarization- and from this, move to the methods used to calculate various physical quantities from the field distributions at the fundamental wavelength, namely spin angular momentum (SAM), super chiral light and SHG polarization.

5.4.1 Linear optical response

Spectral response

As we have seen in the experimental Sec. 5.3, the SHG measurements were achieved using a laser source at 900 nm wavelength. This wavelength was chosen based on the spectral response from the device. In Fig. 5.5, the simulated reflection from the nanostructures for the different gap sizes $d$ is shown. Transmission and
absorption were also calculated, but we are mainly interested in reflection as the SHG measurement is done in this mode. This result clearly shows a peak resonance at around 900 nm which is invariant with the gap size. The reflection peak occurs because of a plasmonic resonance associated with surface currents bound to the metallic nanostructure. As we will see later, when looking at the field distribution at the resonant wavelength, this effect leads to strong/localized field enhancements in the nanostructure which are important for SHG light emission.

Figure 5.5: Modelled reflection spectrum (at normal incidence) from a periodic arrangement of gammadion nanostructures with different gap sizes $d$. Left-CPL was modelled by setting the source to 45° polarization and a 90° phase difference. Numerical convergence was obtained with 12 diffraction orders.

In Fig. 5.6, we show the transmission, reflection and absorption at the resonant wavelength ($\lambda=900$ nm) and varying gap sizes $d$ of the nanostructures. As we have seen with the spectral data in Fig. 5.5, the shape of the reflection curves does not change substantially with the different gap sizes, only their amplitude. This information is better represented in Fig. 5.6, where we observe a slight decrease of reflection with increasing gap size $d$. 
Electromagnetic field distribution

Near-field distribution at the fundamental frequency and normal incidence was investigated using a commercial implementation of the RCWA method discussed in Sec. 3.3.1. In these simulations, numerical convergence of the fields was obtained with 26 diffraction orders for each transverse dimension. The model was built by considering the nanostructure as the main unit cell, and the fields were recorded on a 3D matrix with $513 \times 513 \times 129$ grid points for $x$, $y$ and $z$. Also two simulations were run to obtain the fields for $\theta = 0^\circ$ and $90^\circ$ linear source polarizations, and from these we computed the fields for left- and right-CPL using the following relation [34]

$$ U_i^{\left(\theta\right)}(r) = U_i(r, \theta = 0^\circ) - iU_i(r, \theta = 90^\circ) \quad (5.1) $$

right CPL:  $$ U_i^{\left(\theta\right)}(r) = U_i(r, \theta = 0^\circ) + iU_i(r, \theta = 90^\circ) \quad (5.2) $$

where $U$ is either the electric $E_i$ or magnetic $H_i$ fields, $i$ is one of the three spatial Cartesian components, $\theta$ the incident polarization, and $\mp$ or $\pm$ correspond to right- or left- CPL polarizations.

Figure 5.6: Modelled spectral response for CPL excitation at 900 nm wavelength, as function of the gap size of the $\mathcal{F}_1$ nanostructure. Numerical convergence was obtained with 12 diffraction orders.
Figure 5.7: Spatial profile of the absolute electric field $|\mathbf{E}|$ at the top surface Gold/Air interface of different gap size nanostructures. Top row shows electric field profile on nanostructure excited with right-CPL, and bottom row shows nanostructure excited with left-CPL. Numerical convergence was obtained with 26 diffraction orders.
Chapter 5. SHG-CD response from superchiral metasurfaces

Figure 5.8: Spatial profile of the absolute magnetic field $|\mathbf{H}|$ at the top surface Gold/Air interface of different gap size nanostructures. Top row shows nanostructure excited with right-CPL, and bottom row shows nanostructure excited with left-CPL. Numerical convergence was obtained with 26 diffraction orders.
Figures 5.7 and 5.8 show the absolute value of the electric $|E|$ and magnetic $|H|$ fields respectively at the top gold/air interface of the nanostructures for normal light incident with right-CPL (top row) and left-CPL (bottom row). Each column corresponds to a nanostructure with a different gap size $d$. In the first column we show the $\mathcal{P}$ nanostructure with a gap size of $d=493\,\text{nm}$, in a periodic arrangement of 4 unit cells. The other columns show a magnified version of the fields at the centre of the $\mathcal{P}$ nanostructure (top row) or $\mathcal{S}$ nanostructure (bottom row) representative of the area used for measurement of the SHG intensity during experiment.

Figures 5.7 and 5.8 clearly show the formation of EM-field hotspots at the centre of the nanostructure with chirality matching the one of CPL, coinciding with the location of the SHG sources found in the measurement. The spatial profile of the magnetic field (Fig. 5.8) is homogeneous across the different devices, while larger variations of the local fields are seen in the distribution of the electric-field (Fig. 5.7), in particular for the $\mathcal{S}$ nanostructure as the gap $d$ distance is changed. As to be discussed in Sec. 5.4.4, the SHG polarization is strongly related to the electric-field at fundamental frequency, and therefore, it is expected that the measured SHG signal maps regions of strong local electric-field enhancement. This however is not sufficient to justify the SHG-CD response observed in the SHG experiment, as this is highly dependent on the interactions between multipolar contributions. In the following sections we will look into establishing a relationship between the spatial distribution of superchiral light at fundamental frequency and that of SHG, in particular the physical processes that lead to the SHG-CD response.

### 5.4.2 Orbital angular momentum flux

Linearly polarized light of frequency $\omega$, can be described as having a planar wavefront with wavevector $\vec{k}$ parallel to the propagating direction, carrying both energy $\hbar\omega$ and linear momentum $\hbar k$. In the case of circular or elliptical polarized light, we have optical angular momentum which has a spin component (an intrinsic process associated with the spin of single photons) and an orbital component (an
intrinsic or extrinsic process associated with the helical wavefront [190]). In CPL light, each photon carries $\pm \hbar$ spin angular momentum (SAM) depending on whether it is right- or left- polarized, while for a helical wavefront (with $\exp(i\ell \varphi)$ phase dependence) each photon carries $\ell \hbar$ intrinsic optical angular momentum (OAM) where $\ell \in \mathbb{Z}$. Even though SAM and OAM are independent physical processes, interaction between spin and orbital angular momentum can be achieved using materials with a complex dielectric function, leading to a change in polarization or phase of the light. Various applications for the optical angular momentum of light can be found, including quantum key distribution [191], information encoding [192], and optical particle spanner [193].

A useful quantity for the study of the angular momentum carried by light is the optical angular momentum flux $M_{zz}$, which describes the rate of momentum transfer across an area. In the context of this work, we are particularly interested in looking into the SAM flux ($M_{zz}$) as this is associated with the optical chirality density [194, 195] which we will discuss in Sec. 5.4.3. This quantity can be calculated using the following equation [196]

$$M_{zz}\text{spin}(\omega, \mathbf{r}) = \frac{1}{2\omega} \Re \left[ -i \left( E_x(\omega, \mathbf{r})H_y^*(\omega, \mathbf{r}) + E_y(\omega, \mathbf{r})H_x^*(\omega, \mathbf{r}) \right) \right]$$

where $\omega$ is the incident light frequency, $\mathbf{r}$ the position vector, $\varepsilon_0$ and $\mu_0$ the vacuum permittivity and permeability respectively, $c$ the speed of light in vacuum, $\Re$ the real part of a complex number, and $E$ and $H$ the electric and magnetic fields. Calculation of this quantity for the top air/gold interface of our nanostructures is shown in Fig. 5.9. The top row shows the centre of the \( \mathcal{U} \) unit cell excited with right-CPL, and the bottom row shows the centre of the \( \mathcal{F} \) unit cell excited with left-CPL. The sign of $M_{zz}\text{spin}$ was preserved to distinguish the handedness of the local fields.
Figure 5.9: SAM flux at the gold-air interface of the nanostructures in Fig. 5.2, for normal incidence with right-CPL (top) and left-CPL (bottom) and 900 nm wavelength excitation.
Chapter 5. SHG-CD response from superchiral metasurfaces

Near-field momentum transfer in these nanostructures is dependent on the amplitude and phase gradients associated with the plasmonic fields. Since the EM fields (Fig. 5.7 and 5.8) and SAM flux (Fig. 5.9) are localized at the centre of the unit cell with rotational symmetry identical to the CPL, we expect the reflected fields at these locations to experience an enhancement of optical chirality. For the unit cell and excitation with right-CPL, the magnitude and shape of the SAM flux is invariant with the gap size $d$. This is not the case for the nanostructure, adding to the argument that the variation in SHG-CD found in the experiment is due to non-local effects (i.e. magnetic or higher order multipoles). Note that the SAM flux is an indirect measure of non-local effects due to the Poynting vector ($\mathbf{S} = \mathbf{E} \times \mathbf{H}$) relation between the electric and magnetic fields.

5.4.3 Optical chirality

In this section, we examine chiral optical effects in the nanostructures shown in Fig. 5.2. Our aim is to determine the level of chiral asymmetry generated by the nanostructures to help us establish a connection with the CD response found in the SHG measurements.

An optical EM field incident on a material changes the charge distribution of the atoms or molecules, and the spin of electrons or nuclei. These in turn leads to an electric polarization and magnetization of the material which can act as new sources of radiation. In an achiral medium, radiation effects are mainly dominated by electric-dipole transitions, but this is not the case for a chiral molecule, and electric-magnetic dipole interferences will have a larger role in the optical response of the molecule. Furthermore, higher order multipole transitions will need to be considered if the size of the molecule is comparable to the wavelength of light [197, 198].

When CPL interacts with a chiral molecule, interesting chiral optical effects can be observed since photons experience a change of the complex refractive index as they travel through the medium. This effect leads to (1) optical rotatory dispersion, due to a change in velocity, and (2) circular dichroism, due to a change in light absorption. These effects are usually very weak in biological molecules but can
be greatly improved with certain metallic nanostructures, due to field enhancements from SPP excitations. One way of measuring the enhancement in chiral metasurfaces is by calculation of the optical chirality $C$, which relates the complex electric $E$ and magnetic $H$ field amplitudes using the expression

$$C(\omega, r) = -\frac{\epsilon_0 \mu_0 \omega}{2} \Im [E^*(\omega, r) \cdot H(\omega, r)]$$ (5.4)

where $\omega$ is the incident frequency, $r$ the vector position, $\epsilon_0$ and $\mu_0$ the vacuum permittivity and permeability respectively, and $\Im$ the imaginary part of the dot product. To calculate the local enhancement of the optical chirality $\hat{C}$ we have used the following expression

$$\hat{C}(\omega, r) = \frac{C_s(\omega, r)}{|C_b(\omega, r)|}$$ (5.5)

where $C_s$ and $C_b$ are calculated from Eq. 5.4 with and without the nanostructure respectively. Figure 5.10 shows the spatial distribution of the optical chirality for light excitation with right- and left-CPL (top and bottom respectively). These results show that regions with enhanced optical chirality are localized at the centre of the gammadion nanostructure with the same handedness as CPL. In addition, the location of the enhancement is preserved for both light excitations, with only the sign changing depending on the incident polarization. Similarly to what we have seen with the SAM calculations, we observe invariant changes of amplitude of the enhancement in optical chirality in the $\Gamma\Delta\Upsilon$ nanostructure when excited with right-CPL. In contrast, this is not the case for the $\mathcal{F}$ nanostructure where the amplitude varies with the gap distance. In comparison to identical calculations on a gammadion nanostructure (no gap separation) [189], our results show a more localized enhancement of chirality and also a higher dissymmetry response depending on light polarization.
Figure 5.10: Optical chirality enhancement $|\tilde{C}(r)| > 1$ at the top gold/air interface for the (top) $\text{H}$ nanostructure with right-CPL incidence and (bottom) $\text{H}$ nanostructure with left-CPL incidence. Regions of enhanced optical chirality are located at the centre of the gammadiom shaped structure.
Chapter 5. SHG-CD response from superchiral metasurfaces

Next, we quantify the level of dissymmetry in the chiral optical response of our nanostructures. For CPL incidence, the dissymmetry factor on a chiral molecule can be calculated using the absorption rates for left-CPL ($A^\downarrow$) and right-CPL ($A^\uparrow$), using the following expression [198]

$$g = \frac{A^\downarrow - A^\uparrow}{\frac{1}{2} (A^\downarrow + A^\uparrow)}.$$  \hspace{1cm} (5.6)

The absorption rate ($A^{\uparrow,\downarrow}$) of a molecule to CPL is dependent on two molecular parameters, namely the electric $\alpha$ and electro-magnetic $\beta$ dipole polarizabilities, and also the chirality $C$ and electric energy density $U_e = \epsilon_0 |E|^2 / 2$ of the incident fields. This relationship is expressed as follows [198]:

$$A^{\uparrow,\downarrow} = \frac{2}{\epsilon_0} (\omega U_e \alpha \mp \beta C).$$  \hspace{1cm} (5.7)

This relationship relies on having information on the $\alpha$ and $\beta$ molecular parameters. A more generic definition can be obtained if we consider two chiral enantiomers ($\uparrow$ and $\downarrow$) and accept that the relations $C^{\downarrow}_\uparrow(r) = -C^{\uparrow}_\downarrow(-r)$ and $U^{\downarrow}_e\uparrow(r) = U^{\uparrow}_e\downarrow(-r)$ are true. From these and using Eq. 5.6 and 5.7, we can obtain the following expression for the dissymmetry factor [189]

$$g^* = \frac{C^\downarrow - C^\uparrow}{U^{\downarrow}_e + U^{\uparrow}_e}.$$  \hspace{1cm} (5.8)

Figure 5.11, shows a top view of the enhanced electric energy density $\hat{U}_e r$ at the gold/air interface of the nanostructures for light incident with (top) right-CPL and (bottom) left-CPL. This quantity was calculated using a formula similar to Eq. 5.5 by considering the electric energy density with and without the nanostructure. These results shows that the electric energy density is mainly localized at the corners formed by the L-shaped nanostructures. For the case of the $\uparrow$ nanostructure with right-CPL incidence, the electric energy density maintains more or less the same magnitude for different gap sizes, with slight variations due to numerical imprecisions introduced by RCWA, as discussed in Sec. 5.4.1. In the case of the enantiomer with left-CPL, we observe larger fluctuations of the electric energy density due to constructive destructive interference of the electric fields as a result of the change in gap distance for this structure.
Finally, we compute the enhanced dissymmetry factor \( \hat{g}^* \) for our nanostructures using Eq. 5.8 and similar normalization procedure as in Eq. 5.5. This is shown in Fig. 5.12. In this scenario, only electric dipole-magnetic dipole transitions are considered [69]. For regions with matching strong optical chirality and electric energy density, since \( U_e \) is generally of higher amplitude than \( C \), we will have a dissymmetry that is lower than that of CPL (i.e. \( |g^*| < \frac{2C}{\omega U_e} \)). To obtain a higher dissymmetry (or enantioselectivity), we need low electric energy density and strong optical chirality in the same spatial location. In principle, this can be achieved with these nanostructures by tuning the gap size \( d \), through destructive interference of the electric-field density. According to these calculations, an enhancement of the dissymmetry factor of up to 8 is possible, with distribution in regions around the L-shaped nanostructures.

An additional contribution to the dissymmetry factor is possible through electric dipole-electric quadrupole transitions. This contribution is proportional to \( \Theta_{ijk} = E_i \frac{\delta_{ij}}{5m^2} E_k \) [198], and is similar in formulation to the (nonlocal) bulk SHG polarization from isotropic and centrosymmetric media given by Eq. 2.30. The results for this will be presented in the following section. In summary, we have found that the L-shaped nanostructures are associated with regions of enhanced optical chirality at the centre of the gammadion shapes, and that the enantioselectivity can be tuned by modification of the gap size.
Figure 5.11: Enhanced electric energy density $\hat{\mathcal{E}}(\rho)$ at the top gold/air interface of the (top) $\uparrow\uparrow$ nanostructure with right-CPL and (bottom) $\uparrow\downarrow$ nanostructure with left-CPL.
Figure 5.12: Dissymmetry factor $\hat{g}^*(r)$ at the top gold/air interface of the (top) $\mathcal{H}$ and (bottom) $\mathcal{E}$ nanostructures.
5.4.4 Nonlinear optical response

In this section we will run a numerical analysis of the second-order nonlinear response from the chiral nanostructures in Fig. 5.2. Gold has a face centered cubic structure, which puts it in the class of centrosymmetric materials. As discussed in Sec. 2.2, even-order nonlinear processes are not allowed in centrosymmetric media due to the electric dipole approximation, meaning that electric dipole allowed (local) bulk SHG effects inside the gold nanostructures are not possible. We have however discussed in Sec. 2.2.2 that the symmetry can be broken at the surface of a centrosymmetric material, leading to an electric dipole allowed surface SHG response \( P_s(2\omega, r) \) described by Eq. 2.30. Another SHG source term comes from magnetic dipole and electric quadrupole interactions inside the bulk of centrosymmetric material. This is known as the (nonlocal) bulk SHG response \( P_b(2\omega, r) \) described by Eq. 2.32. The total SHG response of isotropic and centrosymmetric media is then described by the sum of the two components, i.e. \( P(2\omega, r) = P_s(2\omega, r) + P_b(2\omega, r) \).

Regarding the model parameters for the SHG polarization, we have taken the surface nonlinear susceptibilities of gold to be \( \chi_{\|\perp}^{(2)} = \chi_{\perp\|}^{(2)} = 4.63 \times 10^{-20} \text{ m}^2\text{V}^{-1} \), and \( \chi_{\perp\perp\perp}^{(2)} = 1.59 \times 10^{-18} \text{ m}^2\text{V}^{-1} \) [178]. Considering that the fundamental field distribution in the L-shaped nanostructures is highly inhomogeneous, we have used the material parameters \( \beta = 1.29 \times 10^{-32} \), \( \delta = 3.85 \times 10^{-31} \) and \( \gamma = 5.67 \times 10^{-31} \) (all in units of \( \text{A}^3\text{s}^7\text{kg}^{-1}\text{m}^{-3} \)) in the calculation of the (nonlocal) bulk SHG polarization. These were obtained from [123, 199]

\[
\begin{align*}
\beta &= \frac{\varepsilon_0 \varepsilon_r}{8\pi m^* \omega^2} \\
\delta &= -\varepsilon_0 \chi(2\omega) \left( \frac{\chi(\omega)}{e n_e} + \frac{m_e \omega^2 \chi(2\omega)}{e^2 n_e^2} \right) \\
\gamma &= \varepsilon_0 \chi(2\omega) \frac{\chi(\omega)}{2e n_e}
\end{align*}
\]

(5.9a) \hspace{1cm} (5.9b) \hspace{1cm} (5.9c)

where \( m^* = 1.10 m_e \) is the effective mass of gold [200], \( \chi(\omega_g) = \varepsilon_r(\omega_g) - 1, \varepsilon_r(\omega_g) \) the relative permittivity of gold at frequency \( \omega_g \) (calculated using the Lorentz-Drude model from Eq. 2.7), and \( n_e \) the free electron density.
Chapter 5. SHG-CD response from superchiral metasurfaces

Figure 5.13: Spatial profile of the total SHG polarization $\mathbf{P}(2\omega, \mathbf{r})$ for (top) right-CPL and (bottom) right-CPL excitation on the (top) \( \mathcal{U} \) and (bottom) \( \mathcal{H} \) nanostructures with different gap sizes $d$. SHG profile was generated by finding the maximum polarization across the $z$-axis in the gold film. For comparison purposes between devices, the results were normalized with respect to a maximum polarization value.

The total SHG polarization can then be obtained by combining the surface and bulk (nonlocal) SHG contributions from the gold. The spatial profile of the SHG polarization is shown in Fig. 5.13, corresponding to the maximum values obtained across the $z$-axis inside the gold metasurfaces. The results show high localization of the SHG polarization at the centre of the gammadion shape with the same orientation as the incident CPL in agreement with experiment. When comparing these results with the superchiral fields in Fig. 5.10, we also find a matching of regions of strong SHG polarization with areas of superchiral light. A final observation from Fig. 5.13 is that a larger variation of the SHG polarization is seen for the \( \mathcal{H} \) nanostructure, corresponding to the unit cell with varying central distance $d$, whereas for the \( \mathcal{U} \) nanostructure, the SHG profile is consistent across devices.

As we reduce the size of the device to a dimension smaller than the wavelength of light, higher order multipoles beyond the electric dipole approximation will start having a major contribution to the nonlinear response. The width and minimum centre gap of the L-shape metallic structures are 200 nm and 64 nm respectively which is below the 900 nm excitation wavelength. A multipole expansion of the
Chapter 5. SHG-CD response from superchiral metasurfaces

Figure 5.14: SHG power from the (a) \( \mathcal{H} \) nanostructure using left-CPL excitation and (b) \( \mathcal{P} \) nanostructure with right-CPL. The colored data points correspond to (black) electric-dipole, (red) magnetic dipole and (green) electric quadrupole transitions. Yellow area corresponds to the total SHG power of all contributions.

Figure 5.14, shows the total SHG power and individual contributions from the (a) \( \mathcal{H} \) and (b) \( \mathcal{P} \) nanostructures excited with CPL of the same handedness. The results suggest a strong contribution of an electric-quadrupole for both nanostructures, while the electric and magnetic dipoles are 1–2 orders of magnitude lower. In the case of the \( \mathcal{H} \) nanostructure, we observe larger variations of the total SHG power between unit cells, whereas the \( \mathcal{P} \) structure as a more stable emission independent on the gap size \( d \). This result is in agreement with experimental observation of CD effects in the SHG response of the nanostructure. The origin of the SHG-CD we believe is associated with the strong electric quadrupole from the (nonlocal) bulk term of the SHG polarization, owing to the fact that the dissym-
metry factor has a contribution similar in formulation and given by $\Theta_{ijk} = E_i \nabla_j E_k$ [69]. We have however found some difficulties in making a quantitative analysis of these results with the experiment, mainly due to the Gibbs phenomenon originated from the calculation of the local fundamental fields, despite using a large number of harmonics in the model. In the view to improve convergence of the results, it would be of interest to investigate the use of the finite element method in the calculation of the local fields at fundamental frequency. This is known to improve the calculation of the fields around metallic nanostructures subject to rapid spatial variations, as it allows the allocation of a variable mesh with different degrees of resolution.

5.5 Conclusions

In this chapter, we have shown a chiral plasmonic nanostructure with enhanced chiral optical and nonlinear response. From the experimental data, we have seen control over the SHG-CD by change of the gap distance of the $\mathcal{F}$ nanostructure. The numerical studies of the SAM and optical chirality have revealed strong EM interactions and regions of superchiral light at the centre of the gammadion structure with the same handedness as the incident CPL. Additionally, we have shown that the nonlinear response of the device is governed mainly by electric quadrupole contributions, and that control of these transitions is possible by tuning of the gap size of the $\mathcal{F}$ geometry. From the numerical results, we were also able to map regions of superchiral light with regions of large SHG polarization, which suggests that SHG can be used to probe regions of superchiral light, as it is highly sensitive to field enhancements.

A practical application for this type of metasurfaces is in the detection of chiral molecules in a dispersed solution. With the appropriate design to match a particular optical resonance, a large linear response from chiral molecules near the centre of the metasurfaces is to be expected, in comparison to that of CPL with an unstructured background. Nonlinear chiral optical effects in molecules
are also highly sensitive to field enhancements than their linear counterpart, allowing a better contrast in the imaging of molecules of a specific enantiomer. A potential extension to this work could be also to replace the gold film with a hybrid gold-ferromagnet structure, allowing the control over the plasmonic resonance with an external magnetic field [201]. Ferromagnetic materials exhibit magnetic-optical activity, but find limited applications in plasmonics due to the large ohmic losses. However, by integrating this in an alloy mixture with a standard plasmonic material, one can create a system that supports propagating SPP modes with a magneto-tunable response. In principle, one could simply use a fixed metasurface geometry made of a hybrid gold-ferromagnetic material and an external field to tune the plasmonic response to match that of a particular molecule, and therefore enhance the image contrast. Other application would be in optical switching by using a magnetic signal as the control input. Another advantage of hybrid magneto-plasmonic systems is that the magnetic field induced magnetization $\mathbf{M}$ can add a new source term in the description of the surface SHG polarization from the plasmonic structures [202]. This is known as magnetization-induced SHG and is described by $\mathbf{P}_m(2\omega) = \chi_m^{(2)} \mathbf{E}(\omega) \mathbf{E}(\omega) \mathbf{M}$, where $\chi_m^{(2)}$ is a tensor component associated with the magnetization of the material, and has the potential to enhance SHG effects with the added benefit of $\textit{in situ}$ tunability of the nonlinear process.
Chapter 6

Conclusions

In recent years, the field of plasmonics has gained huge momentum and attraction in the scientific community, leading to the development of novel metamaterials with enhanced optical properties that are not typically observed in naturally occurring materials. The ability to confine and manipulate light in metallic nanostructures is therefore of significant interest for the development of future integrated photonic circuits. The next logical step is to create self-contained systems which rely solely on photon-photon interactions to manipulate and process optical signals. Even though significant achievements have been made in nonlinear optics which are paving the way in functionalizing nonlinear light-matter interactions, there are still some limitations with regards to size and power requirements, preventing these technologies from becoming more widespread. A natural progression is therefore to leverage the field enhancement and confinement from highly resonant plasmonic nanostructures, in the optimization and miniaturization of nonlinear optical processes. Research studies in this thesis endeavour to tackle some of these challenges through numerical and experimental investigation of bulk and surface nonlinear optical effects using local field enhancements from plasmonic nanomaterials.
Enhanced bulk nonlinear effects

We have fabricated samples of complementary asymmetric cross-shaped metasurfaces covered with nonlinear materials using EBL, argon-milling, electron-beam evaporation and MBE on a non-conductive substrate. These metasurfaces are characterized by sustaining two orthogonal LSP modes, with intensity controlled by the incident electric-field polarization. Devices have been designed with a response in the near-IR range, and tunable resonances at wavelengths around 1850 nm and 2700 nm. These have been validated via transmission and reflection measurements using FTIR and via numerical calculations of the transmitted power, using a multipole expansion of the linear polarization inside the nonlinear medium. In the linear regime and at normal incidence, electric-dipole transitions have been identified as the dominant process. In addition, experimental and numerical results show that complementary metasurfaces exhibit both a complementary spectral response and field distribution at the resonant frequency, as predicted by Babinet’s principle. However, a question arises regarding the most efficient geometry for enhancing nonlinear optical effects.

ZnO was chosen as the nonlinear dielectric medium, and was epitaxially grown using MBE on top of the complementary metasurfaces, to a thickness of 75 nm. As a non-centrosymmetric material, ZnO exhibits not only second-order but also third-order nonlinear optical effects, which along with the high melting point and wide transparency range, makes it a desirable material for nonlinear optical applications. FTIR measurements were performed before and after the ZnO growth, and a red-shift of the LSP resonances was observed due to an increase of the relative permittivity of the dielectric medium.

We have measured range of nonlinear optical measurements on the complementary nanostructures. SHG and THG nonlinear processes have been only observed in samples containing both the ZnO film and the metasurfaces (there was no nonlinear activity from reference samples), strongly suggesting that these are pure bulk nonlinear processes generated from the ZnO. Due the low birefringence of ZnO, and
small film thickness below the coherence length, the enhanced nonlinear activity is not a phase-matching effect, but the result of strong electric-field localization around the nanostructures, derived from LSP mode excitations. The quadratic and cubic dependence of the SHG and THG signals on the incident power has been investigated, giving further evidence that these are pure nonlinear processes. From the two complementary samples, stronger nonlinear enhancements were observed from the device with transmissive metallic apertures—3.3× and 16.2× larger than the reflective sample, for SHG and THG signals respectively.

The tunable nonlinear optical response from the complementary metasurfaces has also been investigated, both experimentally and numerically, through change of the incident electric-field polarization and/or wavelength. At resonant wavelength, we have observed on average a 11.25dB attenuation of the nonlinear signals as we switch between on- and off- resonant polarizations. We have also observed strong wavelength dependence, with a slight blue-shift of the SHG and THG peaks with respect to the LSP resonance, as result of the relation. Multipole expansion of the nonlinear polarizations has revealed that nonlinear optical processes are characterized by electric-dipole transitions at normal incidence.

Prior to our work with ZnO, we have developed a negative lift-off process allowing to produce arbitrary plasmonic nanoapertures filled with an alternative nonlinear material (Sec. 3.1.7). This process relied on the nanostructuring of a substrate with known nonlinear properties, rather than depositing a nonlinear film at the last stage of fabrication. Some of the key results from this research are presented in Appendix ???. In this work we have used an intrinsic (100) GaAs substrate, also a non-centrosymmetric material with a large \(\chi^{(2)}\) nonlinear susceptibility. Due to the lack of birefringence, the number of applications of GaAs in nonlinear optics is limited, and as a result it can also benefit from the plasmonic enhancements. A range of devices (slit, bowtie and cross-shaped geometries) with a tunable optical response (4 to 12\(\mu\)m range) were produced, and the linear response successfully characterized using the FTIR. Owing to the LSP resonances of these devices being limited to the mid-IR range, we have found that the MCT detectors
Chapter 6. Conclusions

were unsuitable to measure nonlinear signals, despite offering the best sensitivity of all detectors in this frequency range. The high dielectric constant of GaAs and constraints in fabricating devices of reduced dimensions, were the key factors in reaching resonances within the detectivity of more efficient detectors. Further investigation is required towards alternative fabrication techniques to improve the miniaturization of devices and blueshift of resonances.

SHG from chiral metasurfaces

Surface SHG can occur in most materials due to the break in centrosymmetry at the surface. In comparison to (local) bulk effects, these are orders of magnitude lower in amplitude, given the small number of atoms involved in the process. Another contribution comes from (nonlocal) bulk nonlinear effects originated from magnetic dipole and electric quadrupole transitions. In metallic nanostructures, the local field distributions can be highly inhomogeneous, which is highly desirable in the excitation of (nonlocal) bulk effects. Also, the plasmonic enhancements can improve surface-SHG effects in metals by the formation of hot-spots of electric-field radiation. The work in Chap. 5 involved the nonlinear optical characterization of chiral metasurfaces. Our aim was to identify if nonlinear chiral metamaterials can offer enhancements greater than their linear counterparts. We set a control experiment by designing a metasurface composed of two interspersed chiral structures of opposite chirality, and by fabricating variations of this design where the centre of one of the chiral metasurfaces is changed. We observed controllable dichroic SHG response by tuning the central distance of the chiral metasurface, and enhanced SHG intensity for a specific unit cell arrangement. Using a combination of analytical and numerical methods, we were was able to map regions of superchiral light with areas of strong nonlinear polarization. A multipolar analysis of the nonlinear polarization provided further analysis on the dominant contributions to the SHG, in particular the interplay between electric-dipole, magnetic-dipole and electric-quadrupole interactions as we change the centre distance of one unit cell. This work has direct applications in chiral molecular sensing, owing to the large contrast of SHG processes, and in chemical separation, where enhanced lasing fields
could be used to remove a specific enantiomer.

6.1 Future work

In this section, we will conclude with a discussion of some the applications and scientific questions that have been derived from this work.

**Bulk nonlinear optical activity in cross-shaped metasurfaces**

In Sec. 4.5.5, we have initiated experimental studies on the polarization state of the linear transmitted and reflected light, from complementary asymmetric-cross shaped nanostructures. Preliminary results suggest a phase shift of the transmitted/reflected fields at off-resonant polarization, which is in agreement with previous reported studies [96, 203]. This is largely due to the fact that the LSP modes on each arm of the cross can resonant independently and with different phases. Waveplates have already been produced based on these findings, allowing control of the optical activity of linearly polarized light [203], or conversion between linear and circular polarized light [96]. The nonlinear polarization has a complex relationship to the local fields in nonlinear material, and therefore, it would be useful to learn about the polarization state of the nonlinear signals, and the effect of having multiple LSP in the enhancement. These findings could help design more efficient nanostructures with bulk nonlinear optical activity, with applications in light intensity and polarization modulation for integrated photonic circuits [204, 71].

**Laser damage threshold of ZnO films**

Ablation effects and localized heating of the metal have been observed on GaAs filled metallic apertures, due to exposure of the sample to laser intensities as high as $125\text{TW/m}^2$. In the nonlinear measurement of ZnO samples we found no evidence of ablation effects. This could be explained by a combination of the high melting temperature of the ZnO, allowing a better diffusion of the heat across the sample,
and the lower laser intensities used in experiment (1.1TW/m$^2$). Further investigation is required to validate the effectiveness of the thermal properties of the ZnO film.

**Experimental study at off-normal incidence**

At normal incidence, the LSP resonances developed around the asymmetric cross-shaped nanostructures have $E_x(\omega)$ and $E_y(\omega)$ as the dominant components of the electric field, with values an order of magnitude higher than the incident field and dependent on the incoming polarization [90]. From the decomposition of the nonlinear polarizations for ZnO, we have seen that the larger tensor components $\chi^{(2)}_{33}$ and $\chi^{(3)}_{33}$ are largely unexplored at normal incidence (Eq. 4.1b and Eq. 4.1c). There is a potential for significant enhancements of the nonlinear processes at off-normal incidence. In the linear regime, cross shaped nanostructures are known for significant angular stability at transverse magnetic (TM) polarizations [95]. It would be of scientific interest to run a study on the robustness of the nonlinear signals enhanced with the angle of incidence, with applications in active color pixel using an IR source [169].

**Nonlinear optical characterization of home grown ZnO films, alternative growth techniques, and structures**

MBE is certainly a powerful technique to grow single crystal thin films of a variety of materials. However, it does not come without its shortfalls when integrated with other processes involving already deposited materials, influencing the growth conditions across the sample. This problem has been highlighted with the MBE growth of ZnO we performed on the complementary samples, which due to the dominant gold or Al$_2$O$_3$ interfaces, resulted in films of different crystallinities. The nonlinear properties of ZnO have already been reported elsewhere [111], and they can vary significantly depending on the process type, growth conditions, substrate orientation and lattice constant, and even thickness of the film. For an accurate modelling and optimization of bulk nonlinear effects, it would be useful to have a methodic study and measurement of the nonlinear tensor components of films,
based on a range of experimental conditions and device requirements. Although we have seen larger enhancements from plasmonic nanoapertures, an awareness of the nonlinear susceptibility values between complementary samples would have allowed us to make a fairer comparison.

ZnO can also be grown using RF magnetron sputtering, but with reported nonlinear susceptibility values lower than MBE [205, 206]. In principle sputtering is a simpler and more reliable technique, and it can be adapted to a range of substrates with larger deposition rates and reduced running costs. The London Centre for Nanotechnology, boasts a Kurt J. Lesker PVD75 with RF magnetron sputtering sources, which could serve as a testbed for comparison of the physical and optical properties of the ZnO films with the MBE system.

Another great advantage of MBE is the possibility to growth electronic structures of various dimensionalities such as 2DEG, nanowires and quantum dots. For nanowires significant enhancements can be obtained in the nonlinear processes in comparison to bulk effects [111], largely due to the high surface to volume ratio leading to strong surface dipole contributions. ZnO nanowire growth can be achieved through gold seeded nucleation using MBE [207], allowing precise control over the nanowire location and easy integration with plasmonic nanostructures. Quantum dots have also just started being adopted as a gain medium for plasmonic systems [208], opening up new avenues in integrated photonic systems.

Enhanced nonlinear effects from hybrid plasmonic-waveguide mode excitations

Waveguides find applications in nonlinear optics [209]. Owing to their reduced dimensionality, strong EM-field confinement is possible with large propagation lengths and high conversion efficiencies, having practical use in low power systems. In contrast to dielectric waveguides, plasmonic based systems are not limited by diffraction and are capable of an even larger confinement of light in subwavelength dimensions, with applications in highly dense optical interconnects. A major limitation of plas-
monic waveguides is however the large propagation losses in metals, owing to their complex relative permittivity. Recent developments in hybrid optical systems allows the coupling of SPP and dielectric waveguide modes, with large improvements in the propagation of SPPs and stronger field confinements [210]. A logical extension to these developments would be to integrate a nonlinear material as the dielectric medium. The large difference in relative permittivity between ZnO and Al$_2$O$_3$ substrates allows the realization of a whole range of waveguide structures, making ZnO a promising material in hybrid plasmonic-waveguide systems. Furthermore, recent advancements in EM computational methods, incorporating both linear and nonlinear source terms, can now be used to accurately model systems involving a larger number of interacting waves [211], allowing the precise design of photonic structures with enhanced nonlinear activity.
Publications


Conferences and seminars


Bibliography


