Surface modes at metallic and photonic crystal interfaces

by

Weitao Dai

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Condensed Matter Physics

Program of Study Committee:
Costas M. Soukoulis, Major Professor
Rana Biswas
Jiming Song
Ruslan Prozorov
Kerry Whisnant

Iowa State University

Ames, Iowa

2009

Copyright © Weitao Dai, 2009. All rights reserved.

TABLE OF CONTENTS

LIST OF FIGURES v

ACKNOWLEDGEMENTS xiv

CHAPTER 1. Introduction 1

CHAPTER 2. Fundamentals 5

2.1 Surface plasmons 5

2.1.1 Surface plasmons at an interface 5

2.1.2 Surface plasmons of multilayer systems 9

2.1.3 Localized surface plasmons around subwavelength metal particles 13

2.2 Surface waves at photonic crystal surfaces 16

2.2.1 Introduction to photonic crystals 16

2.2.2 Band structures of photonic crystals 17

2.2.3 Surface modes at photonic crystal surfaces 20

2.2.4 Surface modes along a layer of dielectric rods 23

2.3 Extraordinary transmission and beaming 23

2.3.1 Extraordinary transmission 23

2.3.2 Beaming 26

CHAPTER 3. Surface Waves Along a Metal-Dielectric Interface with Defects 29

3.1 Background 29

3.2 Surface impedance boundary conditions 31

3.2.1 Validity of SIBC 32

3.2.2 Surface plasmons based on SIBC 34

3.2.3 Waveguide modes based SIBC 34
3.2.4 Summary of SIBC 36
3.3 Green's function 36
3.3.1 Surface wave at a metal-dielectric interface with defects 36
3.3.2 Simplification of the Green's function 38
3.3.3 Discussion 41
3.3.4 Conclusion 45
CHAPTER 4. Theory of Extraordinary Transmission 46
4.1 Background 46
4.2 Mode expansion method 46
4.3 Transmission into a 1D subwavelength grating with infinite length 51
4.3.1 Simplification of the general equations 51
4.3.2 The PEC grating 53
4.3.3 The SIBC grating 54
4.4 Transmission through a 1D subwavelength grating with finite length 58
4.4.1 The input surface 58
4.4.2 The output surface 60
4.4.3 Transmission through a finite-length grating 61
4.4.4 Effect of hole shapes 66
4.4.5 Oblique incidence 67
4.4.6 Conclusion 69
4.5 Application of the extraordinary transmission theory 69
4.5.1 Phys. Rev. Lett. **83**, 2845 (1999) by J. A. Porto *el al.* 70
4.5.2 Phys. Rev. Lett. **92**, 183901 (2004) by K. J. Klein Koerkamp *el al.* 72
4.5.3 Phys. Rev. Lett. **96**, 233901 (2006) by Z. Ruan and M. Qiu 74
4.5.4 Conclusion 77
CHAPTER 5. Beaming 78
5.1 Introduction of beaming 78

5.2 Beaming of two-layer dielectric rods 78
5.2.1 A two-layer structure 78
5.2.2 More two-layer structures 86
5.2.3 Conclusion 88
5.3 Control of beaming angles 90
5.3.1 Oblique beaming 90
5.3.2 Introduction of a metallic beaming structure 90
5.3.3 Control of beaming angles 92
5.3.4 Frequency splitter 98
5.3.5 Conclusion 100
APPENDIX A. Simulations of infinitely long waveguides 101
A.1 Perfect matched layer 101
A.2 Perfect electric conductor waveguide 104
A.3 Metallic waveguide 106
A.4 Photonic waveguide 107
BIBLIOGRAPHY 113
AUTHOR PUBLICATIONS 120

LIST OF FIGURES

Figure 2.1 Schematic representation of a flat metal-dielectric interface. The interface is $y = 0$ and the fields are TM polarized. 6

Figure 2.2 Dispersion relation of the surface plasmons at a silver-air interface. The black line is the light line. The silver is described by the Drude model. 1 PHz = 10^{15} Hz. 7

Figure 2.3 H_z distributions of the surface plasmon when $\epsilon_m = -2$ and $\epsilon_d = 1$. 8

Figure 2.4 Schematic representation of a symmetrical three layer structure. The middle layer has relative permittivity ϵ_1 the upper and lower layers have the relative permittivity ϵ_2. The thickness of the middle layer is W. The fields are TM polarized. 9

Figure 2.5 Dispersion relation of the coupled surface plasmon mode for an air-silver-air multilayer structure. The thickness of the silver layer is 50 nm. The dispersion relation of a single silver-air interface is plotted too with the label "Single interface". The black line is the light line. The silver is described by the Drude model with negligible damping. 1 PHz = 10^{15} Hz. 11

Figure 2.6 Dispersion relation of the basic mode of an silver-air-silver multilayer structure. The dispersion relation of a single silver-air interface is plotted too with the label "Single interface". The black line is the light line. The silver is described by the Drude model with negligible damping. 1 PHz = 10^{15} Hz. . . 12

Figure 2.7 Schematic representation of a homogeneous subwavelength sphere illuminated by an incident plane wave. The sphere has radius R and relative permittivity ϵ_m surrounded by a homogeneous, lossless medium with relative permittivity ϵ_d. 13

Figure 2.8 Figure excerpted from the paper J. Comput. Theor. Naonosci. **5**, 2096 (2008) by L. Hu *et al.*. The field distributions around the silver sphere with radius $a = 10$ nm embedded in silicon. The field distributions are normalized to the incident plane wave with $\lambda = 800$ nm, which is the surface plasmon wavelength of silver spheres embedded in silicon. 15

Figure 2.9 Figure excerpted from the book ”Photonic crystals: Molding the Flow of Light” by J. D. Joannopoulos *et al.*. Simple examples of photonic crystals. The different colors represent materials with different refractive indices. . . . 16

Figure 2.10 The photonic band structure of a 2D photonic crystal. (a) the cross-sectional view of the photonic crystal in one unit cell. The photonic crystal is a square array of square alumina rods embedded in air with side length $d = 3.1$ mm and $\epsilon = 9.8$. The lattice constant is $a = 11$ mm. (b) The Brillouin zone. (c) The photonic band structure. The blue lines represent TM modes and the red represent TM modes. 19

Figure 2.11 Band structures based on the supercell method. The top figures in (abcd) show the calculation area. The photonic crystal is the sames as the one in Fig. 2.10. The lattice constant is $a = 11$ mm. The bottom figures are the band structures. The gray stripe shows the band gap. In (bcd), the supercell contains 20 photonic crystal layers, only 6 are shown in the diagram. In (c), the length of the air layers at both ends of the photonic crystal is $10a$. In (d) the length of the air layers is $9a$ and circular rods with radius $r = 1.83$ mm and relative permittivity $\epsilon = 9.8$ are added at both ends of the photonic crystal. The black line is the light line. 21

Figure 2.12 Band structure of one layer of dielectric rods. The top figure shows the calculation area with size of 220 mm$\times$11 mm. The rods have circular cross section with radius r = 1.83 mm and relative permittivity ϵ = 9.8. The bottom shows the band structures based on the supercell method. The black line is the light line. 22

Figure 2.13 Figure excerpted from the paper Nature **445**, 39 (2007) by C. Genet and T. W. Ebbesen. The normalized transmission spectrum through a 225 nm-thick Au film with a triangle lattice of circular holes under normal incidence. The hole diameter is 170 nm and the lattice constant is 520 nm. Before and after the film are the glass substrate and index-matching liquid. η is the normalized transmission. 24

Figure 2.14 Figure excerpted from the paper Opt. Lett. **26**,1972 (2001) by T Thio *et al.*. The normalized transmission spectrum through a Ag hole surrounded by (A) rings and (B) a square lattice of dimples. The rings have sinusoidal cross section. The mean radii is $R_k = kP$, $P = 750\ nm$ and $k = 1, 2, \cdots, 10$. The peak-to-peak amplitude is h. The dimples have lattice constant 750 nm and depth h = 180 nm. 25

Figure 2.15 Figure excerpted from the paper Phys. Rev. B **76**, 235417 (2007) by R. Moussa *et al.* partially. Left column shows the exit surface of a photonic crystal. From top to bottom: a bare photonic crystal, a photonic crystal with a surface layer, a photonic crystal with a surface layer and grating layer. The right column shows the field distribution after the waveguide. 27

Figure 3.1 Figure excerpted from the paper Nature Phys. **2**, 551 (2006) by P. Lalanne and J. P. Hugonin. The transmission through the subwavelength metallic slit accompanied by a grooves gotten from experiments, numerical simulations and analytic calculations. Blue circles: experimental data. Solid blue line: analytic results based on CDEW model. Solid black line: numerical data computed by the fully vectorial modal method. Red dots: analytic results based on surface plasmon mode. The Inset: the geometry of the metallic slit-groove structure. 30

Figure 3.2 Schematic representation of light incidence at a metal-dielectric interface under (a) TE polarization and (b) TM polarization. 32

Figure 3.3 Schematic representation of a two-dimensional metallic waveguide. 35

Figure 3.4 Schematic representation of a dielectric-metal interface with indentations in a two-dimensional space. 37

Figure 3.5 Integration path of the non-dimensional Green's function in the complex space of h. The red curves are the integration path and the blue lines are the branch cut. 39

Figure 3.6 $|g_C(s)/g_{SP}(s)|$ along a metal-air interface when the metal permittivity is $\epsilon_m = -31.39 + 2.22i$. 41

Figure 3.7 Real and imaginary parts of H_z along a silver-air interface in the regions $y \in (0.1\lambda, 10\lambda)$ and $y \in (20\lambda, 30\lambda)$ when $\lambda = 3000$ nm and $\epsilon_m = -329+47.5i$. The simulation results are calculated by COMSOL Multiphysics, a commercial finite element method software. The other curves are calculated analytically using different simplified forms of the Green's function shown in the legends. 43

Figure 3.8 Real and imaginary parts of H_z along a metal-air interface in the regions $y \in (0.1\lambda, 10\lambda)$ and $y \in (30\lambda, 40\lambda)$ when $\lambda = 500$ nm and $\epsilon_m = -8.50 + 6.00i$. The simulation results are calculated by COMSOL Multiphysics. The other curves are calculated analytically using different simplified forms of the Green's function shown in the legends. 44

Figure 4.1 Schematic picture to study the transmission into infinitely-long slits. 47

Figure 4.2 The normalized transmission into a PEC subwavelength grating with infinite length illuminated by a normal incident plane wave. The incident wavelength is $\lambda = 560$ nm and the slit width of the grating is $W = 30$ nm. 55

Figure 4.3 The normalized transmission into a SIBC subwavelength grating with infinite length illuminated by a normal incident plane wave. The incident wavelength is $\lambda = 560$ nm, the relative permittivity of the metal is $\epsilon_m = -7.92$, and the slit width of the grating is $W = 30$ nm. The upper and lower figures have the same simulation curve. In the upper curve, the analytic curve is calculated using the approximation of the Green's function $g(s) \approx g_{SP}(s) + g_C(s)$. The analytic curve in the lower figure is calculated using the approximation $g(s) \approx g_{SP}(s)$. 57

Figure 4.4 Schematic picture to of a metallic grating with a finite length. 58

Figure 4.5 The maximum transmission a PEC grating can reach by changing the grating length at different grating periods, The slit width is $W = 30$ nm and the incident wavelength is $\lambda = 560$ nm. 62

Figure 4.6 The analytic maximum transmission a PEC grating can reach by changing the grating length. The incident wavelength is $\lambda = 560$ nm. 63

Figure 4.7 The COMSOL Multiphysics simulation results of the maximum transmission a SIBC grating can reach by changing the grating length at different grating periods, The slit width is $W = 30$ nm and the incident wavelength is $\lambda = 560$ nm. The relative permittivity of the metal is $\epsilon_m = -7.92$ 64

Figure 4.8 Analytic transmission through a PEC grating with a finite length.The incident wavelength is $\lambda = 560$ nm. 66

Figure 4.9 Figures excerpted from PRL **83**, 2845 (1999). (Top) Schematic representation of the metallic grating studied in the paper. (Below) Zero-order transmission through the grating vacuum at normal incidence for different values of the grating length. The grating period is $d = 3.5\ \mu$m and the slit width is $a = 0.5\ \mu$m. 71

Figure 4.10 Figures excerpted from PRL **92**, 183901 (2004). (Top) Images of two periodic hole arrays used. Both arrays have period $d = 425$ nm. (a) Circular holes with a diameter of 190 nm (b) Rectangular holes of 75 nm $\times$ 225 nm. (Below) Calculated normalized transmission of three different subwavelength periodic hole arrays. E field of the incident wave was perpendicular to the long axis of the rectangles. 73

Figure 4.11 A figure excerpted from PRL **96**, 233901 (2006). Schematic representation of a PEC film with a periodic square array consisting of rectangular air holes studied in the paper. 75

Figure 4.12 Figures from excerpted PRL **96**, 233901 (2006). (Top) Normalized transmission through a PEC film with periodic hole arrays for different lattice constants d. The thickness of the film is $0.2a$. The size of the hole is $0.9a \times 0.2a$. (Below) The mean value of the normalized transmission of the random sample with different cell size. The solid purple line shows the normalized transmission of the periodic hole array with the lattice constant $d = 1.0a$. The hole size of the random sample is still $0.9a \times 0.2a$. The filling ratio of the holes equal to that of the periodic array. The inset shows a example of a sample with 25 randomly distributed holes in a $5a \times 5a$ cell. 76

Figure 5.1 (a) A schematic drawing of the two-layer structure and (b) a schematic drawing of the simulation area. 79

Figure 5.2 Simulation results of empty area (a, b), a surface layer only (c, d), a grating layer only (e, f) and a two-layer structure (g, h). The left column shows how the transmission and the distribution factor change with frequency. The right column shows the electric field along the red line right after the rods when $f = 11.59$ GHz. 81

Figure 5.3 E_z distribution of the simulation area (a) without and (b) with a two-layer structure when $f = 11.59$ GHz. 82

Figure 5.4 Distribution factors of fields after the two-layer structure illuminated by by oblique Gaussian beams. ϕ is the incident angle. 83

Figure 5.5 Band structure of the surface layer along with the wavevectors of the surface modes which can be excited by the incident waves when the incident angles are $0°, 5°$ and $10°$. 84

Figure 5.6 Transmission of oblique Gaussian beams through the two-layer structure. . . 85

Figure 5.7 Transmission and distribution factor of the field after two two-layer structures. The black line shows the fitted transmission using the equation of the transmission matrix theory. 87

Figure 5.8 Time average energy density distribution when the simulation area has 2,4,6 and 8 two-layer structure. The frequency is $f = 11.59$ GHz and the distance between two adjacent structure is $d = 21.29\lambda$. 89

Figure 5.9 Schematic representation of the subwavelength metal slit surrounded by upper and lower periodic grooves. 91

Figure 5.10 (a)&(b) Amplitudes and phases of fields at symmetric grooves with $a = b =$ 500 nm; (c)&(d) the linear fitting results of phases of the fields at grooves for different a when $b = 500$ nm and for different b when $a = 500$ nm. 93

Figure 5.11 (a) Normalized far-field angular transmission when $a = b = 500$ nm; (b) Constructive interference angles at different groove periods; (c)&(d) Normalized far-field angular transmission when $b = 446$ nm and $b = 710$ nm. 95

Figure 5.12 (a)&(b) Amplitudes and phases of fields at grooves above the slit when the grooves are symmetric with $a_1 = a_2 = 920$ nm, $b_1 = b_2 = 446$ nm and when there are only upper periodic grooves with $a_1 = 920$ nm, $b_1 = 446$ nm; (c)&(d) Normalized far-field angular transmission generate by the slit with upper periodic grooves alone when $b_1 = 446$ nm and $b_1 = 710$ nm. 96

Figure 5.13 (a) Normalized far-field angular transmission of a slit with upper periodic grooves alone when $a_1 = 980$ nm, $b_1 = 446$ nm and a slit with lower periodic grooves alone with $a_2 = 860$ nm, $b_2 = 710$ nm; (b) Normalized far-field angular transmission of a slit with asymmetric upper and lower periodic grooves when $a_1 = 980$ nm, $b_1 = 446$ nm, $a_2 = 860$ nm, $b_2 = 710$ nm. 97

Figure 5.14 (a) Heights and (b) full-widths at half-maximum of the highest far-field transmission peak at different angles generated by upper periodic grooves only and asymmetric upper and lower periodic grooves. 98

Figure 5.15 Normalized far-field angular transmission at two frequencies when (a) $a_1 = 600$ nm, $b_1 = 500$ nm, $a_2 = 400$ nm, $b_2 = 500$ nm and (b) $a_1 = 540$ nm, $b_1 = 450$ nm, $a_2 = 360$ nm, $b_2 = 450$ nm. The groove depth is $h = 500$ nm. . 99

Figure A.1 Schematic representation of incidence from isotropic dielectric in $x < 0$ half-plane to perfect matched layer in $x > 0$ half-plane. 102

Figure A.2 Schematic representation of a UPML in a PEC or a SIBC waveguide. 105

Figure A.3 Schematic representation of a two-dimensional photonic crystal waveguide with infinite length. The left region is an ordinary photonic crystal (PC) waveguide; the right blue region is the PC-based uniaxial perfectly matched layer (UPML). $x = 0$ is the PC-UPML interface. 107

Figure A.4 Schematic representation of a two-dimensional photonic crystal grating with infinite length. The left region is the ordinary photonic crystal (PC) grating; the right blue region is the PC-based uniaxial perfectly matched layer (UPML). $x = 0$ is the PC-UPML interface. The incident wave is TM polarized. 109

Figure A.5 Transmsission into a photonic crystal grating followed by free space, a homogeneous UPML and a PC-based UPML. The homogeneous UPML has $\bar{\bar{\epsilon}}_{rod} = \bar{\bar{\epsilon}}_{air} = \epsilon_0 \bar{\bar{s}}$ and $s_x = 1+0.3i$. The PC-based UPML has $N = 20$, $M = 0.3$, $p = 0$. The total number of layers of the graing changes from 20 to 40. 110

Figure A.6 Transmsission into a photonic crystal grating followed by PC-based UPMLs. The parameters of the PC-based UPMLs are $N = 20, M = 0.3, p = 1$ and $N = 20, M = 0.3, p = 2$. The total number of layers of the graing changes from 20 to 40. 111

ACKNOWLEDGEMENTS

First and foremost I offer my sincerest gratitude to my advisor: Dr. Costas M. Soukoulis for his encouragement and guidance during my graduate career. He supported me with his patience, knowledge and funds while allowing me the room to work in my own way. He taught me how to be a good physicist. I am deeply indebted to him.

I would like to express my gratitude to Prof. Rana Biswas, Prof. Jiming Song, Prof. Ruslan Prozorov and Prof. Kerry Whosnant for serving in my committee. Special thanks to Prof. Rana Biswas, Prof. Jiming Song and Prof. Ruslan Prozorov for the knowledge I learned from their courses. I also want to thank Prof. Jianwei Qiu as a former committee member. I appreciate for their time, interest, and helpful comments about my thesis.

My colleagues in the Iowa State University supported and helped me so much in my research work. Especially, I thank Dr. Thomas Koschny. He is always ready to offer me help and answer me any questions. I benefited greatly from the instructive discussion with him. I would like to thank my office mates Dr. Bingnan Wang, Anan Fang and Rongkuo Zhao for the friendly and vigorous office atmosphere. I also want to thank Dr. Jiangfeng Zhou, Dr. Lei Zhang, Dr. Durdu Güney and Dr. Rabia Moussa for their help.

I would like to take this opportunity to express my thanks to all the persons who offer me help for my education and research. It is impossible to list all these people's names. But my my gratitude is not less than for those listed above.

On the personal side, I would like to thank my parents for raising me and offering me love unconditionally; thank my brother for growing up with me. Finally special thanks to my girlfriend Qian Wang for her love.

CHAPTER 1. Introduction

A surface mode is an electromagnetic field distribution bounded at a surface. It decays exponentially with the distance from the surface on both sides of the surface and propagates at the surface. The surface mode exists at a metal-dielectric interface as surface plasmon (1) or at a photonic crystal surface terminated properly (34; 35; 36). Besides its prominent near-filed properties, it can connect structures at its propagation surface and results in far-field effects. Extraordinary transmission (EOT) and beaming are two examples and they are the subjects I am studying in this thesis.

EOT means the transmission through holes in an opaque screen can be much larger than the geometrical optics limitation. Based on our everyday experience about shadows, the transmission equals the filling ratio of the holes in geometrical optics. The conventional diffraction theory also proved that the transmission through a subwavelength circular hole in an infinitely thin perfect electric conductor (PEC) film converges to zero when the hole's dimension is much smaller than the wavelength (40). Recently it is discovered that the transmission can be much larger than the the filling ratio of the holes at some special wavelengths (41). This cannot be explained by conventional theories, so it is called extraordinary transmission.

It is generally believed that surface plasmons play an important role (43; 44) in the EOT through a periodic subwavelength hole array in a metallic film. The common theories in literatures are based on these arguments. The surface plasmons cannot be excited by incident plane waves directly because of momentum mismatch. The periodicity of the hole arrays will provide addition momentum. When the momentum-matching condition of surface plasmons is satisfied, the surface plasmons will be excited. Then these surface plasmons will collect the energy along the input surface and carry them to the holes. So the transmission can be bigger than the filling ratio. Based on this picture, they deduced naturally that when surface plasmons' momentum-matching condition is satisfied, the surface plasmons

are excited sufficiently and the transmission reaches its peak.

I present a new theory from first principles to explain EOT through one-dimensional periodic subwavelength metallic slits in this thesis. This theory can also be extended to 2D hole arrays. I define the incident wavelengths that satisfy the momentum-matching condition as surface resonant wavelengths. I proved analytically that the transmission is actually zero at the surface resonant wavelengths. The correct logic is: When the momentum-matching condition is satisfied, the surface plasmons excited by each slit interfere constructively with each other, the total surface plasmons will go to infinity. But the law of nature forbids the infinity. The only solution is the surface plasmon excited by one slit is zero and all the energy is reflected.

In my theory, the term corresponding to surface plasmons appear explicitly in the equations. So it confirms the importance of surface plasmons without any doubt. The theory divides the transmission process into two steps: energy collection process along the input surface and the propagation process in the slits. In the first process, the surface plasmons collect the energy along the input surface and carry them to the slits. This process happens efficiently at any wavelength other than the surface resonant wavelengths. So EOT can happen at almost any wavelength. After the energy enter the slits, the Fabry-Pérot interference between the input and output surface decides how much energy is emitted from the slits. So the EOT wavelengths are decided by the Fabry-Pérot resonances.

I also use my theory to explain the data in literatures. The transmission spectra through 1D slits or 2D hole arrays in literatures agree with my theory very well. The new theory can explain a lot of experimental results published recently, such as the transmission through randomized hole arrays, the strong influence of the hole shape on the transmission peaks, and so on.

Beaming is another far-field effect resulting from surface modes. Normally light coming from a subwavelength waveguide is diffracted to all angles. With the help of surface modes, we can confine the output field in a small angle interval. This phenomenon is called beaming (46).

The principle of the beaming has been explained clearly in literatures (47). To achieve good beaming, a photonic crystal waveguide need a surface layer to support surface modes and a grating layer to coupling the evanescent surface modes into propagation modes. A metallic beaming structure is generally a subwavelength waveguide surrounded by periodic structures such as grooves or dielectric

gratings (53; 54). The flat metal surface supports the surface mode, so additional surface layer is not necessary. The periodic structures work as the grating layer.

We discovered that a single layer of dielectric rods can support surface modes (39). Thus I design a very simple beaming structure: two-layer dielectric-rod structure. The first layer is used to support the surface modes and the second couples the surface modes into propagation modes. The photonic crystal waveguide is not a required component of a beaming structure.

Numeric simulations show clearly one two-layer dielectric-rod structure converges the incident Gaussian beam and two structures in series achieve good beaming and full transmission simultaneously. By repeating periodically this two-layer structure one can obtain excellent beaming and enhanced transmission for very long distances. Theoretically this simple beaming structure confirms the importance of surface modes in beaming. In practice, I design a novel waveguide for Gaussian beams.

Beside the research on forward beaming , oblique beaming was also demonstrated using either metallic structures (75; 76) or photonic crystals (77) in literatures. I developed an efficient method to design a structure that can achieve oblique beaming at any angle between 0 and 70 degrees.

The structure is a subwavelength metallic slits surrounded by upper and lower periodic grooves at the output surface. The output surface is equivalent to several point sources. By changing the geometry parameters of the grooves, I can control the fields, specially the phases of the fields, of the point sources, and control the output field distributions furthermore. I find the period of the grooves decides the beaming angle and the distance between the slit and the grooves decides the beaming intensity. By adjusting the geometry parameters of the upper and lower grooves, I can tilt the beam to any angle and have very good beaming quality. I find surprisingly the best beaming happens not at the forward direction but an oblique direction. I also design a frequency splitter: the beaming angles of two close wavelengths deviate with each very much.

The rest chapters in this thesis are organized in this way:

In the second chapter I prove the existence of the surface modes at a metal-dielectric interface or a photonic crystal surface with a surface layer. I also briefly review the EOT phenomenon and beaming phenomenon.

In the third chapter, I calculate the surface waves along a metal-dielectric interface with an inden-

tation analytically. The surface waves is basically surface plasmons with some residual waves. The complete description of the surface wave sets up a solid foundation to understand EOT and beaming.

In the fourth chapter, I present my theory about EOT. In the fifth chapter, I discuss two beaming structures: the two-layer dielectric-rod structure and the metallic oblique beaming structure.

Then in the appendix, I discuss how to simulate the waveguide with infinite length numerically.

CHAPTER 2. Fundamentals

A surface mode is an electromagnetic field distribution bounded at a surface. It decays exponentially with the distance from the surface on both sides of the surface and propagates at the surface. Surface modes exist at metallic surfaces and photonic crystal surfaces. In this thesis, I will study the properties of the two kinds of surface modes and two phenomena strongly connected to surface modes: extraordinary transmission (EOT) and beaming.

2.1 Surface plasmons

Surface plasmons are surface modes trapped at a metallic surface through their interaction with the free electrons of the metal (1). The fundamentals and applications of surface plasmons are nicely reviewed in some books (2).

2.1.1 Surface plasmons at an interface

Over a wide frequency range, the relative permittivity of a metal can be described by the Drude model (3; 4)

$$\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + i\omega_c)}. \tag{2.1}$$

Here ω_c is the damping frequency and

$$\omega_p^2 = \frac{ne^2}{\epsilon_0 m_{eff}}$$

is the plasma frequency, where n is the electron density, e is the electron charge and m_{eff} is the effective optical mass of an electron. Noble metals have extremely high plasma frequencies. For example, sliver has $\omega_p = 2\pi \times 2184$ THz and $\omega_c = 2\pi \times 4.35$ THz (5). When ω is smaller than the plasma frequency, the metal's permittivity has negative real part. Comparing to the electric fields, the interactions between the magnetic fields and the electrons are weak. So the relative permeability $\mu(\omega) \approx 1$.

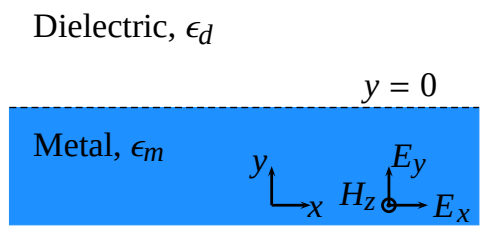

Figure 2.1 Schematic representation of a flat metal-dielectric interface. The interface is $y = 0$ and the fields are TM polarized.

The simplest structure supporting surface plasmons is a flat metal-dielectric interface. Fig. 2.1 shows an interface at $y = 0$. The fields are TM polarized. So the fields in the dielectric region are

$$
\begin{aligned}
H_z(x, y) &= A \exp[i\beta x] \exp[-k_d y]; \\
E_x(x, y) &= -i \frac{k_d}{\omega \epsilon_0 \epsilon_d} A \exp[i\beta x] \exp[-k_d y]; \\
E_y(x, y) &= \frac{\beta}{\omega \epsilon_0 \epsilon_d} A \exp[i\beta x] \exp[-k_d y].
\end{aligned}
$$

and the fields in the metal region are

$$
\begin{aligned}
H_z(x, y) &= A \exp[i\beta x] \exp[k_m y]; \\
E_x(x, y) &= i \frac{k_m}{\omega \epsilon_0 \epsilon_m} A \exp[i\beta x] \exp[-k_d y]; \\
E_y(x, y) &= \frac{\beta}{\omega \epsilon_0 \epsilon_m} A \exp[i\beta x] \exp[-k_d y].
\end{aligned}
$$

Here ϵ_m and ϵ_d are the relative permittivities of the dielectric and metal.

H_z is continuous along the $y = 0$ interface, so the magnitude of the magnetic field A and the x component of the wave vector β are the same in the two regions. Because of the continuity of E_x along the interface, we get

$$
\frac{k_m}{k_d} = -\frac{\epsilon_m}{\epsilon_d}.
$$

The relative permittivities satisfy $\mathrm{Re}(\epsilon_m) < 0$ and $\epsilon_d > 0$. So there is a solution making $\mathrm{Re}(k_m) > 0$ and $\mathrm{Re}(k_d) > 0$, which is the surface plasmon solution.

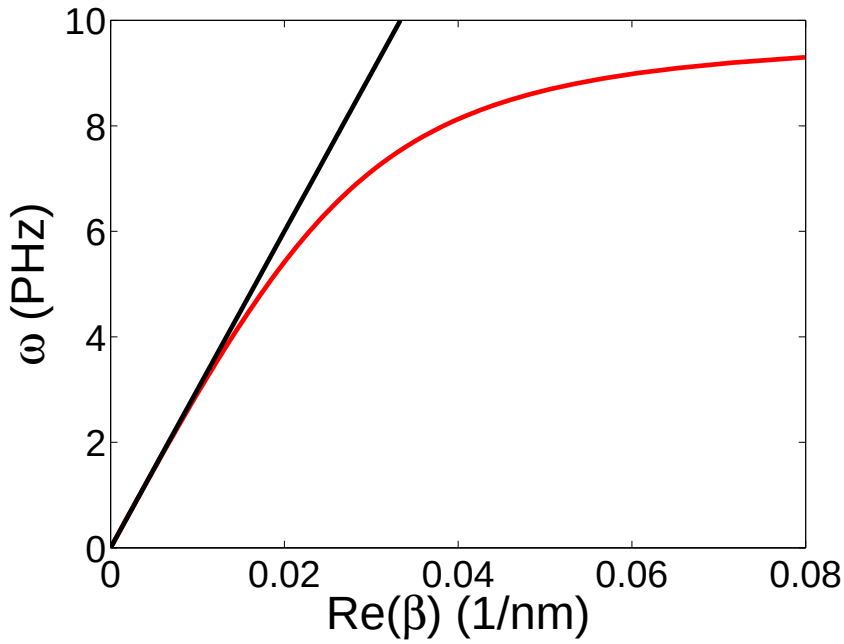

Figure 2.2 Dispersion relation of the surface plasmons at a silver-air interface. The black line is the light line. The silver is described by the Drude model. 1 PHz = 10^{15} Hz.

The fields in the dielectric and metal have to fulfill the dispersion relations of the media, so

$$\frac{\omega^2}{c^2}\epsilon_d = \beta^2 - k_d^2;$$
$$\frac{\omega^2}{c^2}\epsilon_m = \beta^2 - k_m^2.$$

Then we get the dispersion relation of the surface plasmons propagating along the interface

$$\beta = \frac{\omega}{c}\sqrt{\frac{\epsilon_m \epsilon_d}{\epsilon_m + \epsilon_d}}. \tag{2.2}$$

Fig. 2.2 shows the Dispersion relation of the surface plasmons at a silver-air interface, where the relative permittivity of the silver is calculated form Eq. (2.1). H_z distributions of the surface plasmon is plotted in Fig. 2.2 when $\epsilon_m = -2$ and $\epsilon_m = 1$. We can clearly see the fields decay exponentially on the both sides of the interface.

If we neglect the damping frequency in Eq. (2.1), then

$$\epsilon_m(\omega) \approx 1 - \omega_p^2/\omega^2.$$

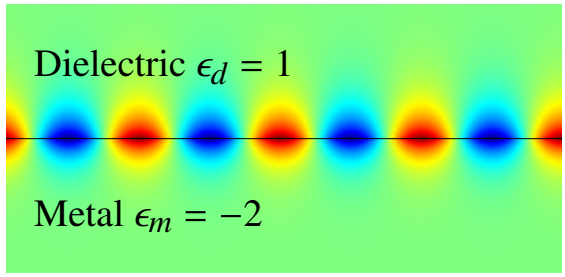

Figure 2.3 H_z distributions of the surface plasmon when $\epsilon_m = -2$ and $\epsilon_d = 1$.

When ω approaches the surface plasmon frequency

$$\omega_{sp} = \sqrt{\frac{\omega_p}{1+\epsilon_d}},$$

$\epsilon_m + \epsilon_d \to 0$ and β goes to infinity. Fig. 2.2 shows the converging property.

Suppose the real and imaginary parts of ϵ_m are ϵ' and ϵ'' respectively and $|\epsilon'| \gg |\epsilon''|$. Then $\beta = \beta' + i\beta''$ with

$$\beta' = \frac{\omega}{c}\left(\frac{\epsilon'\epsilon_d}{\epsilon'+\epsilon_d}\right)^{1/2};$$
$$\beta'' = \frac{\omega}{c}\left(\frac{\epsilon'\epsilon_d}{\epsilon'+\epsilon_d}\right)^{3/2}\frac{\epsilon''}{2(\epsilon')^2}.$$

The surface plasmon can be characterized by three lengths:

1. The propagation length of the surface plasmon $L = (2\beta'')^{-1}$;

2. The skin depth in the dielectric

$$\hat{z}_d = \frac{1}{|k_d|} \approx \frac{\lambda}{2\pi}\left(\frac{\epsilon'+\epsilon_d}{-\epsilon_d^2}\right)^{1/2};$$

3. The skin depth in the metal

$$\hat{z}_m = \frac{1}{|k_m|} \approx \frac{\lambda}{2\pi}\left(\frac{\epsilon'+\epsilon_d}{-(\epsilon')^2}\right)^{1/2};$$

When the frequency is much smaller than the surface plasmon frequency, ϵ' is very negative and β is close to the light line of the dielectric. So $\hat{z}_d$ is big, the surface plasmons extend into the dielectric.

When ω approaches ω_{sp}, $\hat{z}_d \to 0$ and $\hat{z}_m \to 0$. So the energy is confined in a very small area. We can use this property to design surface plasmon waveguides.

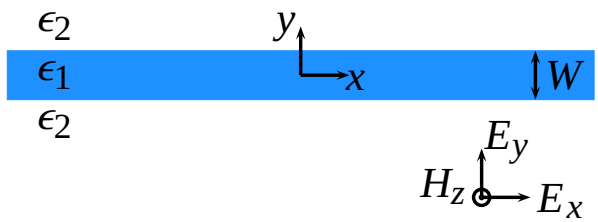

Figure 2.4 Schematic representation of a symmetrical three layer structure. The middle layer has relative permittivity ϵ_1 the upper and lower layers have the relative permittivity ϵ_2. The thickness of the middle layer is W. The fields are TM polarized.

2.1.2 Surface plasmons of multilayer systems

Now I begin to study the the multilayer structures composed by alternative metal and dielectric layers. Every metal-dielectric interface supports surface plasmons. When the layer is very thin, the surface plasmons at neighboring interfaces will interact with each other and bring us new physics (12; 13; 14).

Here I only study the simplest multiple layer structures: symmetrical three layer structures. The top and bottom layers are half-infinite with the same relative permittivity ϵ_2. In the middle is a layer with relative permittivity ϵ_1 and width W. $y = 0$ is the middle line of the middle layer. I only consider TM polarization here.

Given β as the wavevector along x direction, the field in the middle region is

$$H_z = e^{i\beta x}(Ae^{-k_1 y} + Be^{k_1 y});$$
$$k_1 = \sqrt{\beta^2 - \epsilon_1 k_0^2}.$$

Here $k_0 = \omega/c$ is the wavevector in free space.

H_z in the upper and lower layers are

$$H_z = Ce^{i\beta x}e^{-k_2 y} \text{ when } y > W/2;$$
$$H_z = De^{i\beta x}e^{ik_2 y} \text{ when } y < -W/2.$$
$$k_2 = \sqrt{\beta^2 - \epsilon_2 k_0^2},\ \mathrm{Re}(k_2) > 0.$$

H_z and E_x are continuous along the interfaces $x = \pm W/2$. So

$$Ae^{-k_1 W/2} + Be^{k_1 W/2} = Ce^{-k_2 W/2}; \tag{2.3}$$

$$k_1/\epsilon_1(Ae^{-k_1 W/2} - Be^{k_1 W/2}) = k_2/\epsilon_2 Ce^{-k_2 W/2}; \tag{2.4}$$

$$Ae^{k_1 W/2} + Be^{-k_1 W/2} = De^{-k_2 W/2}; \tag{2.5}$$

$$k_1/\epsilon_1(Ae^{k_1 W/2} - Be^{-k_1 W/2}) = -k_2/\epsilon_2 De^{-k_2 W/2}. \tag{2.6}$$

Form Eq. (2.3) and Eq. (2.4), we get

$$\frac{Ae^{-k_1 W/2} + Be^{k_1 W/2}}{Ae^{-k_1 W/2} - Be^{k_1 W/2}} = \frac{\epsilon_2 k_1}{\epsilon_1 k_2}. \tag{2.7}$$

Form Eq. (2.5) and Eq. (2.6), we get

$$-\frac{Ae^{k_1 W/2} + Be^{-k_1 W/2}}{Ae^{k_1 W/2} - Be^{-k_1 W/2}} = \frac{\epsilon_2 k_1}{\epsilon_1 k_2}. \tag{2.8}$$

Put the two equations together, we have

$$\frac{Ae^{-k_1 W/2} + Be^{k_1 W/2}}{Ae^{-k_1 W/2} - Be^{k_1 W/2}} = -\frac{Ae^{k_1 W/2} + Be^{-k_1 W/2}}{Ae^{k_1 W/2} - Be^{-k_1 W/2}}. \tag{2.9}$$

It is easy to verify this equation is equivalent to $A^2 = B^2$. So the dispersion relation is

$$\frac{\epsilon_2 k_1}{\epsilon_1 k_2} = \frac{e^{-k_1 W/2} + e^{k_1 W/2}}{e^{-k_1 W/2} - e^{k_1 W/2}} \quad \text{or} \tag{2.10}$$

$$\frac{\epsilon_2 k_1}{\epsilon_1 k_2} = \frac{e^{-k_1 W/2} - e^{k_1 W/2}}{e^{-k_1 W/2} + e^{k_1 W/2}}. \tag{2.11}$$

Eq. (2.10) describes the odd modes and Eq. (2.11) describes the even modes in the system. When $W \to \infty$, the two equations become

$$\epsilon_2^2 k_1^2 = \epsilon_1^2 k_2^2$$

and the dispersion relation returns to Eq. (2.2). The two interfaces decouple.

Fig. 2.5 shows the dispersion relation of an air-silver-air multilayer structure. ϵ_1 is calculated from a Drude model and $\epsilon_2 = 1$.

The odd modes of the dielectric-metal-dielectric multilayer structure have a special property: when the thickness of metal layer decreases, the coupled surface plasmons tend to stay out of the metal. Since absorption happens in the metal, the odd modes have longer propagation length (15). The even modes have the opposite property: they tend to stay in the metal, which reduce the propagation length.

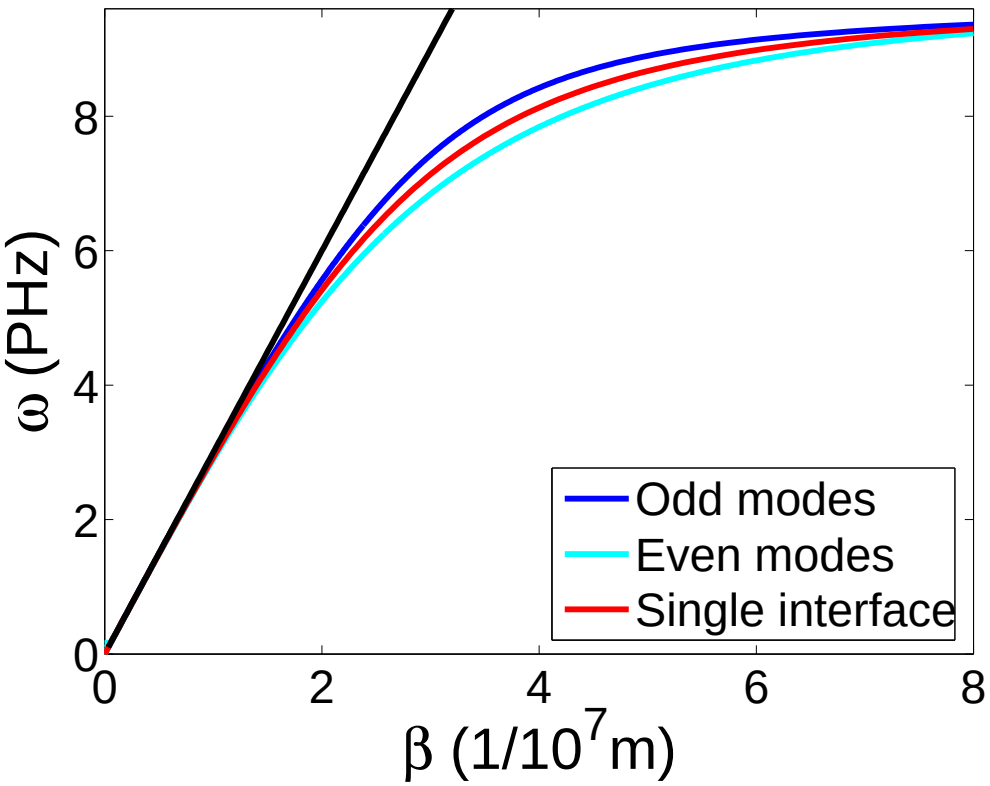

Figure 2.5 Dispersion relation of the coupled surface plasmon mode for an air-silver-air multilayer structure. The thickness of the silver layer is 50 nm. The dispersion relation of a single silver-air interface is plotted too with the label "Single interface". The black line is the light line. The silver is described by the Drude model with negligible damping. 1 PHz = 10^{15} Hz.

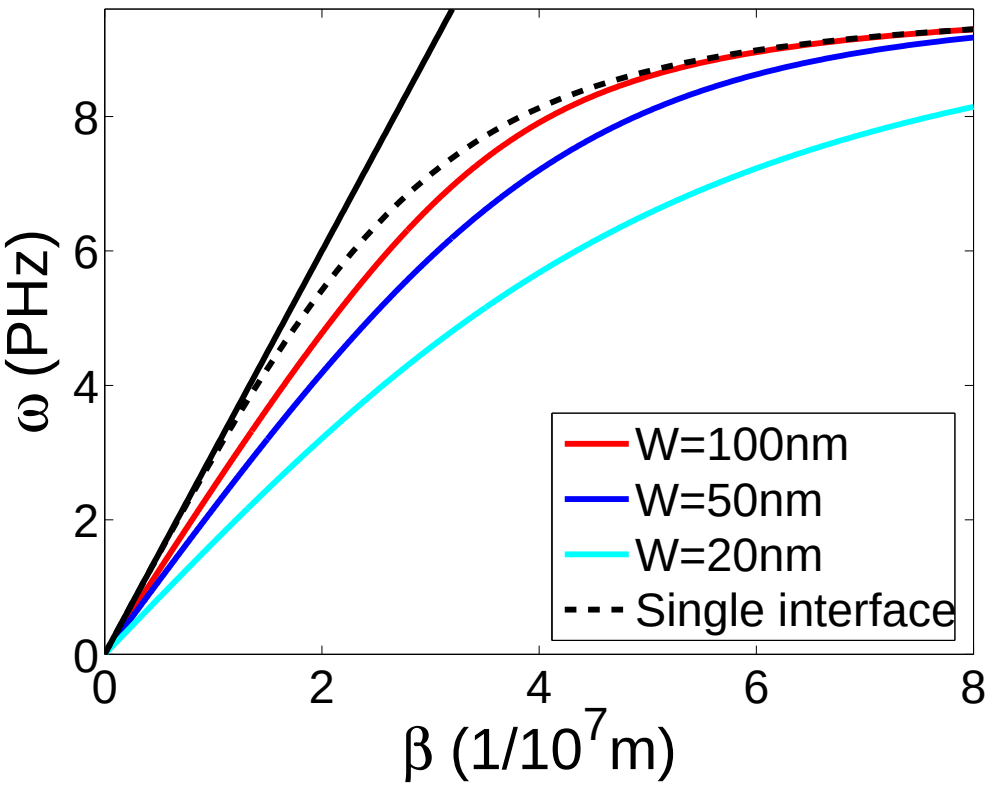

Figure 2.6 Dispersion relation of the basic mode of an silver-air-silver multilayer structure. The dispersion relation of a single silver-air interface is plotted too with the label "Single interface". The black line is the light line. The silver is described by the Drude model with negligible damping. 1 PHz = 10^{15} Hz.

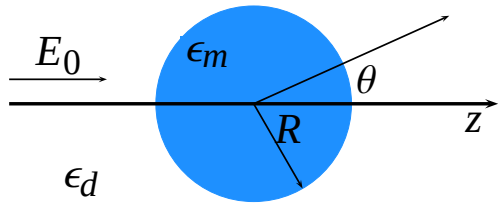

Figure 2.7 Schematic representation of a homogeneous subwavelength sphere illuminated by an incident plane wave. The sphere has radius R and relative permittivity ϵ_m surrounded by a homogeneous, lossless medium with relative permittivity ϵ_d.

I now move to the metal-dielectric-metal multilayer structure. The dispersion relation of the basic mode of an silver-air-silver multilayer structure is shown in Fig. 2.6 for three different thickness of the air layer. The basic mode is interesting because the energy of the basic mode is mostly confined in the air layer. No matter how narrow the dielectric layer is, the basic mode is a propagation mode inside the dielectric layer. We can find that β of the mode changes with W, which give us a possible way to control the fields (16). when W is very small, β is big even at ω much smaller than ω_{sp}, which is impossible for a single interface. Big β insures the energy is confined in a very small region, even much smaller than the wavelength, so the metal-dielectric-metal multilayer structures works perfectly as waveguides (17; 18; 19).

2.1.3 Localized surface plasmons around subwavelength metal particles

Surface plasmons can also be excited at the surfaces of subwavelength metallic particles. The surface plasmons enhance fields around the small particles .

The localized surface plasmons around a subwavelength metal sphere in a plane wave can be calculated analytically. Fig. 2.7 shows the system under study: the sphere has radius R, relative permittivity ϵ_m and is surrounded by a homogeneous, lossless medium with relative permittivity ϵ_d.

When the size of the sphere is much smaller than the incident wavelength, we can neglect the phase oscillation and assume the sphere locates in a electrostatic field. This is called quasi-static approximation (20). Then the Maxwell's equations are simplified to the Laplace equation of the potential $\nabla^2\Phi = 0$ and the electric field is $\vec{E} = -\nabla\Phi$. The external field becomes $\vec{E} = E_0\hat{z}$.

Because of the azimuthal symmetry of the system, the Laplace equation has the general solution is (20)

$$\Phi(r,\theta) = \sum_{l=0}^{\infty}[A_l r^l + B_l r^{-(l+1)}]P_l(\cos\theta).$$

Here $P_l(\cos\theta)$ are the Legendre polynomials, r is the length of the position vector $\vec{r}$ and θ is the zenith angle.

Because the potential is finite at the origin, the potentials inside and outside the sphere are

$$\Phi_{in}(r,\theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos\theta) \qquad \text{when } r \le R; \tag{2.12}$$

$$\Phi_{out}(r,\theta) = \sum_{l=0}^{\infty}[B_l r^l + C_l r^{-(l+1)}]P_l(\cos\theta) \qquad \text{when } r > R. \tag{2.13}$$

A_l, B_l and C_l are decided by the boundary conditions. First when $r \to \infty$, Φ_{out} converges to $-E_0 z$. So $B_l = -E_0\delta_{l1}$ where δ_{l1} is the Kronecker delta. Next considering the continuity of the potential function and the tangential components of the electric field at the surface of the sphere, finally we can get

$$\Phi_{in}(r,\theta) = -\frac{3\epsilon_d}{\epsilon_m + 2\epsilon_d}E_0 r\cos\theta; \tag{2.14}$$

$$\Phi_{out}(r,\theta) = -E_0 r\cos\theta + \frac{\epsilon_m - \epsilon_d}{\epsilon_m + 2\epsilon_d}\frac{R^3}{r^2}E_0\cos\theta. \tag{2.15}$$

The potential outside the sphere is basically the superposition of the external field and the field of a dipole located at the origin. Define the dipole moment

$$\vec{p} = 4\pi\epsilon_0\epsilon_d R^3\frac{\epsilon_m - \epsilon_d}{\epsilon_m + 2\epsilon_d}\vec{E}_0.$$

Then

$$\Phi_{out}(r,\theta) = -E_0 r\cos\theta + \frac{\vec{p}\cdot\vec{r}}{4\pi\epsilon_0\epsilon_d r^3}.$$

Electric fields can be calculated from $\vec{E} = -\nabla\Phi$:

$$\vec{E}_{in} = \frac{3\epsilon_d}{\epsilon_m + 2\epsilon_d}\vec{E}_0;$$
$$\vec{E}_{out} = \vec{E}_0 + \frac{3\vec{n}(\vec{n}\cdot\vec{p}) - \vec{p}}{4\pi\epsilon_0\epsilon_d r^3}$$

We can define the polarizability α as $\vec{p} = \epsilon_0\epsilon_d\alpha\vec{E}_0$, then

$$\alpha = 4\pi R^3\frac{\epsilon_m - \epsilon_d}{\epsilon_m + 2\epsilon_d}. \tag{2.16}$$

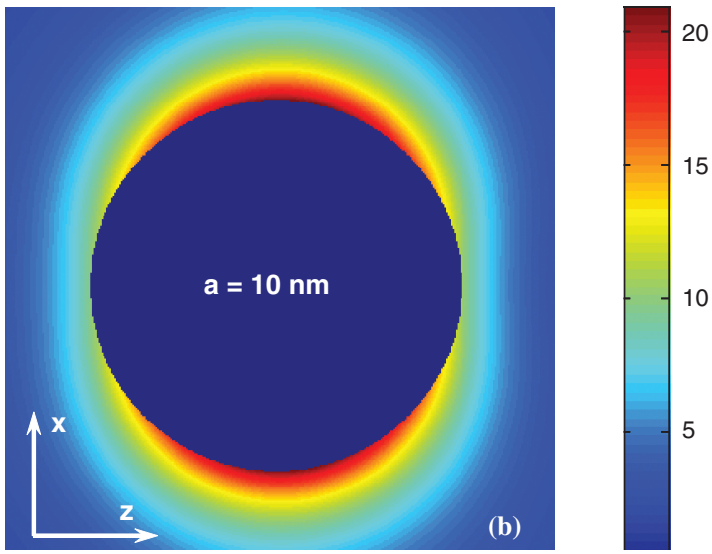

Figure 2.8 Figure excerpted from the paper J. Comput. Theor. Naonosci. **5**, 2096 (2008) by L. Hu *et al.*. The field distributions around the silver sphere with radius $a = 10$ nm embedded in silicon. The field distributions are normalized to the incident plane wave with $\lambda = 800$ nm, which is the surface plasmon wavelength of silver spheres embedded in silicon.

The polarizability experiences a resonance when

$$\mathrm{Re}(\epsilon_m) = -2\epsilon_d. \tag{2.17}$$

This relation is named as the Fröhlich condition and the frequency satisfying this condition is called the surface plasmon frequency of a metal sphere.

Now I will use the electrostatics calculation results above to time-harmonic electromagnetic fields. When the incident wave is $\vec{E}(r,t) = \vec{E}_0 \exp[-i\omega t]$ and $R \ll \lambda$, the dipole representation is sill valid and $\vec{p}(t) = \epsilon_0\epsilon_d\alpha\vec{E}_0 \exp[-i\omega t]$. Then corresponding scattering and absorbing cross sections of a sphere are (21)

$$C_{sca} = \frac{k^4}{6\pi}|\alpha|^2;$$

$$C_{abs} = k\mathrm{Im}(\alpha).$$

So the scattering and absorption of a metallic sphere are strongly enhanced at the surface plasmon frequency.

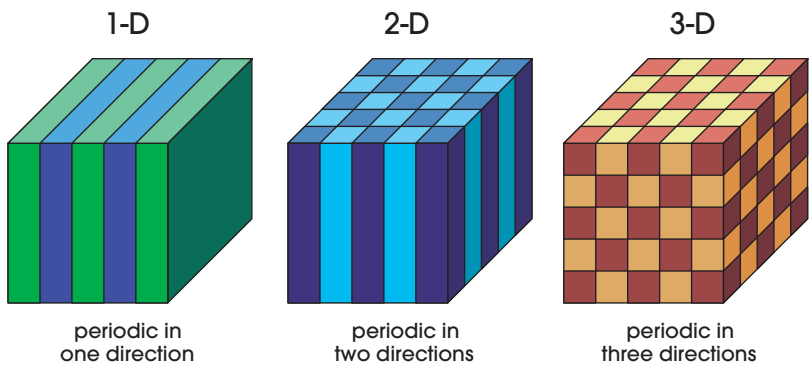

Figure 2.9 Figure excerpted from the book "Photonic crystals: Molding the Flow of Light" by J. D. Joannopoulos *et al.*. Simple examples of photonic crystals. The different colors represent materials with different refractive indices.

Fig. 2.8 shows the the field distributions around the silver sphere with radius $a = 10$ nm when $\lambda = 800$ nm (22). The fields are normalized to the incident plane wave. We can find the fields around the particle are much bigger than the incident fields. These strong localized fields are very useful to generate Raman scattering (23; 24), fluorescence (25) and other nonlinear phenomena (26).

When the particle becomes bigger, the quasi-static approximation fails and we have to use a rigorous electrodynamics approach. The field distribution then follows Mie solution (21; 27).

2.2 Surface waves at photonic crystal surfaces

2.2.1 Introduction to photonic crystals

Photonic crystals provide powerful ways to manipulate electromagnetic fields. Some comprehensive books were published recently addressing this topic (6; 7; 8). Here I just give a simple introduction of photonic crystal.

Simply speaking, photonic crystals are just periodic arrays of materials with different refractive indices (6). Some simple examples of photonic crystals are shown in Fig. 2.9.

When the period of the photonic crystal (or lattice constant) is much bigger than the wavelength, we can neglect the wavelength and use geometrical optics (9). The electromagnetic waves become light rays. When the wavelength is much larger than the lattice constant, we can neglect the details in one

unit cell of the photonic crystal and treat the photonic crystal as a homogeneous medium. Actually the normal homogeneous materials are not homogeneous at the level of atoms.

The distance between two neighboring atoms in a crystal is several angstroms. With the development of nanotechnology, persons can make structures around or even smaller than 100 nm (10). These structures are much larger than atoms, which gives scientists a lot of freedom to design them. These structures are also much smaller than the wavelength of infrared and even visible lights, so they are regard as effective materials. These artificial structures are called metamaterials, which can exhibit exceptional properties not readily observed in nature (11).

Photonic crystals mostly focus on the region between the two limitations where the wavelengths are of the same order of the lattice constants. Both the approximation to rays and the approximation to homogeneous media fail. We have to solve the Maxwell's equations and use band structure to describe photonic crystals.

2.2.2 Band structures of photonic crystals

Similar to the crystals, we can also use the band structure to describe the electromagnetic field distributions supported by photonic crystals.

Suppose the photonic crystal is characterized by the periodic relative permittivity $\epsilon(\vec{r})$ and the relative permeability is $\mu = 1$. $\vec{a}_i$ where $i = 1, 2, 3$ are the primitive vectors of the unit cell of the photonic crystal. The periodicity of $\epsilon(\vec{r})$ implies

$$\epsilon(\vec{r} + \vec{a}_i) = \epsilon(\vec{r}) \quad \text{for } i = 1, 2, 3.$$

From the Maxwell's equations

$$\begin{aligned}
&\nabla \times \vec{E}(\vec{r}, t) = -\mu_0 \frac{\partial}{\partial t} \vec{H}(\vec{r}, t); \\
&\nabla \times \vec{H}(\vec{r}, t) = \epsilon(\vec{r}) \epsilon_0 \frac{\partial}{\partial t} \vec{E}(\vec{r}, t); \\
&\nabla \cdot \left[\epsilon(\vec{r}) \vec{E}(\vec{r}, t)\right] = 0; \\
&\nabla \cdot \vec{H}(\vec{r}, t) = 0,
\end{aligned}$$

We will get

$$\frac{1}{\epsilon(\vec{r})}\nabla\times(\nabla\times\vec{E}(\vec{r},t)) = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\vec{E}(\vec{r},t);$$
$$\nabla\times\left(\frac{1}{\epsilon(\vec{r})}\nabla\times\vec{H}(\vec{r},t)\right) = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\vec{H}(\vec{r},t).$$

Suppose $\vec{E}(\vec{r},t)$ and $\vec{H}(\vec{r},t)$ are the time-harmonic fields with the form $\vec{E}(\vec{r},t) = \vec{E}(\vec{r})\exp[-i\omega t]$ and $\vec{H}(\vec{r},t) = \vec{H}(\vec{r})\exp[-i\omega t]$. Then $\vec{E}(\vec{r})$ and $\vec{H}(\vec{r})$ are the eigenfunctions of the operators $\mathcal{L}_E$ and $\mathcal{L}_H$ defined below:

$$\mathcal{L}_E\vec{E}(\vec{r}) \triangleq \frac{1}{\epsilon(\vec{r})}\nabla\times(\nabla\times\vec{E}(\vec{r})) = \frac{\omega^2}{c^2}\vec{E}(\vec{r}); \tag{2.18}$$

$$\mathcal{L}_H\vec{H}(\vec{r}) \triangleq \nabla\times\left(\frac{1}{\epsilon(\vec{r})}\nabla\times\vec{H}(\vec{r},t)\right) = \frac{\omega^2}{c^2}\vec{H}(\vec{r}). \tag{2.19}$$

Then we can use Bloch's theorem to Eq. (2.18) and Eq. (2.19). The solutions of the two equations are then characterized by a wavevector $\vec{k}$ and a band index n and have the form

$$\vec{E}_{\vec{k}n}(\vec{r}) = u_{\vec{k}n}(\vec{r})\exp[i\vec{k}\cdot\vec{r}];$$
$$\vec{H}_{\vec{k}n}(\vec{r}) = v_{\vec{k}n}(\vec{r})\exp[i\vec{k}\cdot\vec{r}].$$

and $u_{\vec{k}n}(\vec{r})$ and $v_{\vec{k}n}(\vec{r})$ are periodic functions with the same unit cell as $\epsilon(\vec{r})$. The band structures describe the relations between the wavevector $\vec{k}$ and the corresponding eigenfrequency ω.

An example of the photonic band structure is shown in Fig. 2.10. The photonic crystal is a two-dimensional one composed by a square array of square alumina rods embedded in air. The rods are of dimension $d = 3.1$ mm with relative permittivity $\epsilon = 9.8$. The lattice constant is $a = 11$ mm.

One important property of the band structure is the existence of band gap under TM polarization (electric field parallel with the rods). No eigenmode exists when the frequency is between 9.54 Hz and 12.81 GHz. So the waves with these frequencies can not propagate in the photonic crystal. The complete photonic band gap for any propagation directions and any polarizations has been found with 3D photonic crystals (28; 29).

Photonic crystals offer us the opportunities to "design" the band structures. A lot of applications come from the unusual band structures of photonic crystals. For example, the band gap leads to the photonic crystal waveguides directly (30; 31). Other applications includes superprisms (32) and superlens (33).

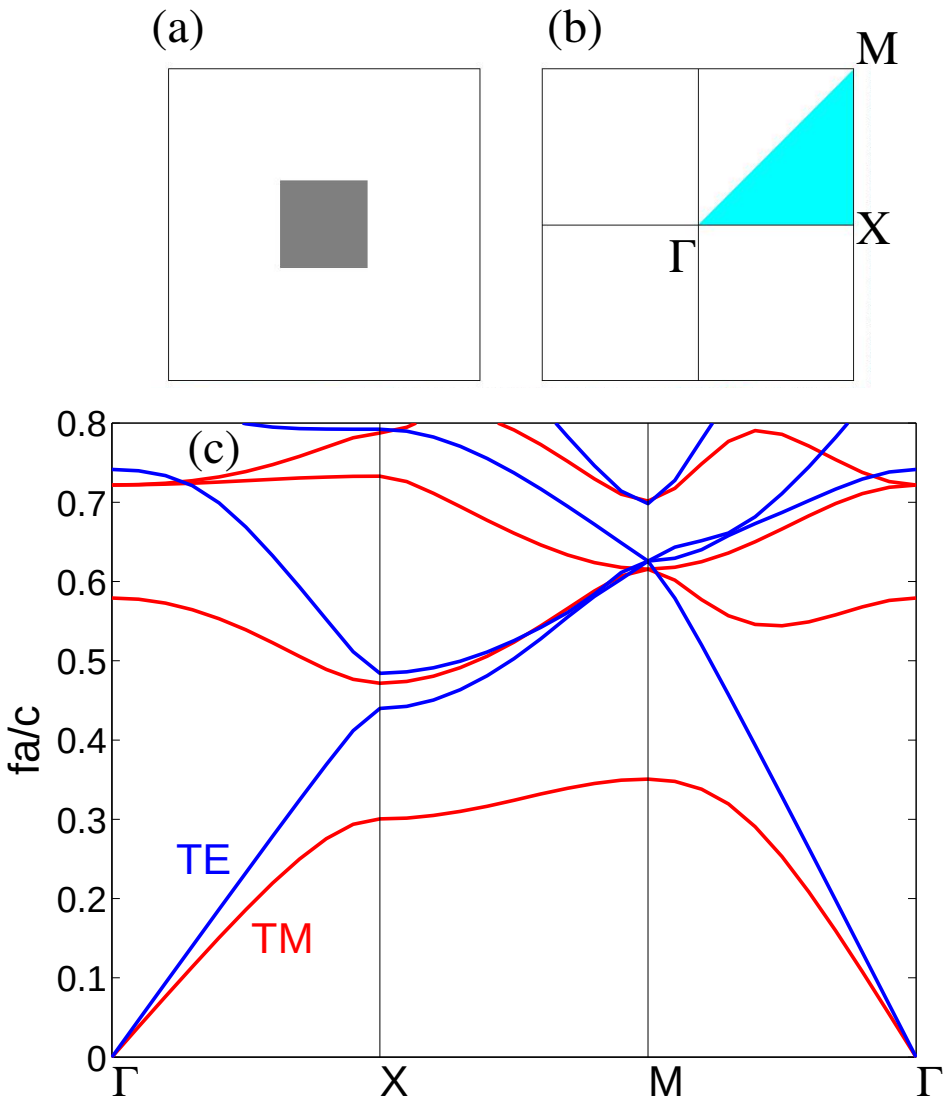

Figure 2.10 The photonic band structure of a 2D photonic crystal. (a) the cross–sectional view of the photonic crystal in one unit cell. The photonic crystal is a square array of square alumina rods embedded in air with side length d = 3.1 mm and ϵ = 9.8 . The lattice constant is a = 11 mm. (b) The Brillouin zone. (c) The photonic band structure. The blue lines represent TM modes and the red represent TM modes.

2.2.3 Surface modes at photonic crystal surfaces

The existence of the surface modes at photonic crystal surfaces has been verified theoretically (34; 35; 36) and experimentally(37; 38; 39). It was found that the surface modes are sensitive to the surface termination and the bear photonic crystal surface don't support surface modes. Further research is necessary to understand this behavior.

To study surface modes, we can employ that supercell method. Supercell is a cell containing more than one unit cells of a photonic crystal. The band structures based on the supercell method are shown in Fig. 2.11.

The photonic crystal used here is the same what used in Fig. 2.10. Fig. 2.11(a) shows the band structure calculated based on the unit cell of the photonic crystal. This band structure is equivalent the band structure in Fig. 2.10. I calculate the eigenfrequencies for different (k_x, k_y) but only show the corresponding k_y in the figure. The gray stripe in the band structure figure shows the band gap. In Fig. 2.11(b), I calculate the eigenfrequencies based on the supercell. The supercell contains 20 unit cell of the photonic crystal. The diagram shows only 6 of them for simplicity. I set that the fields along the left boundary equal the fields along the right boundary; the fields along the bottom boundary has a phase difference $\exp[ik_y a]$ with the fields along the top boundary (a is the lattice constant). Since the supercell has 20 unit cells along x direction, $\exp[ik_x \times 20a] = 1$ means $\exp[ik_x \times a]$ can have 20 different values. So one time eigenfrequency calculation based on supercell is equivalent to 20 times calculations using different k_x values based on a unit cell. So I get the same band structure in Fig. 2.11(b) as in Fig. 2.11(a).

In Fig. 2.11(c) I add an air layer between photonic crystals. The supercell contains 20 unit cells of the photonic crystals in the middle (only 6 are shown in the diagram of Fig. 2.11(c) and air layers with length $10a$ at both ends. Repeating the supercell along x and y directions, we can not get a perfect photonic crystal but the alternative air and photonic crystal layers. So the band structure here is different from the band structure of the perfect photonic crystal shown in Fig. 2.11(a) and Fig. 2.11(b). For example some bands exist in the band gap of the photonic crystal. These modes will stay in the air layers.

In Fig. 2.11(d) I add surface layers at both ends of the photonic crystal. The surface layers is

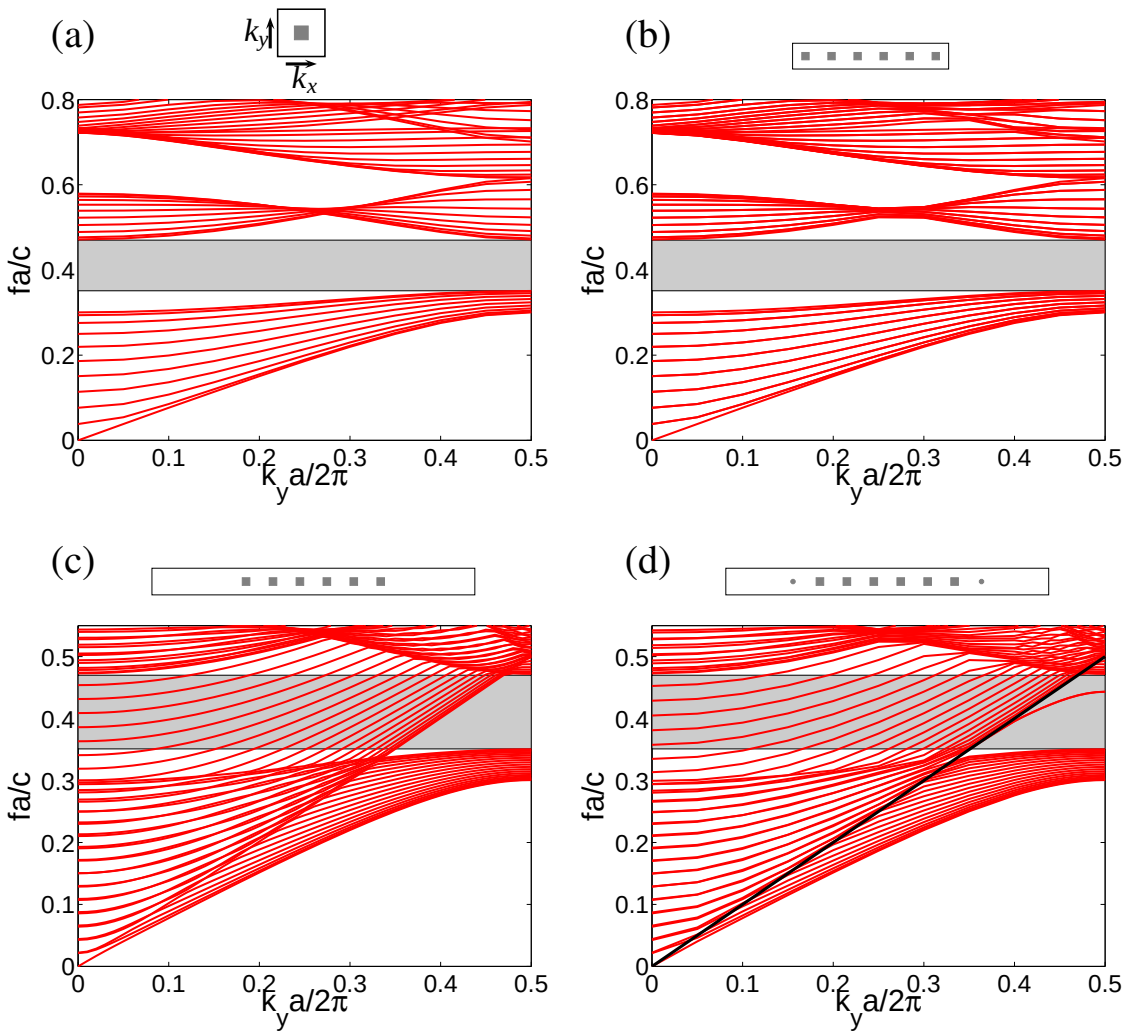

Figure 2.11 Band structures based on the supercell method. The top figures in (abcd) show the calculation area. The photonic crystal is the sames as the one in Fig. 2.10. The lattice constant is $a = 11$ mm. The bottom figures are the band structures. The gray stripe shows the band gap. In (bcd), the supercell contains 20 photonic crystal layers, only 6 are shown in the diagram. In (c), the length of the air layers at both ends of the photonic crystal is $10a$. In (d) the length of the air layers is $9a$ and circular rods with radius $r = 1.83$ mm and relative permittivity $\epsilon = 9.8$ are added at both ends of the photonic crystal. The black line is the light line.

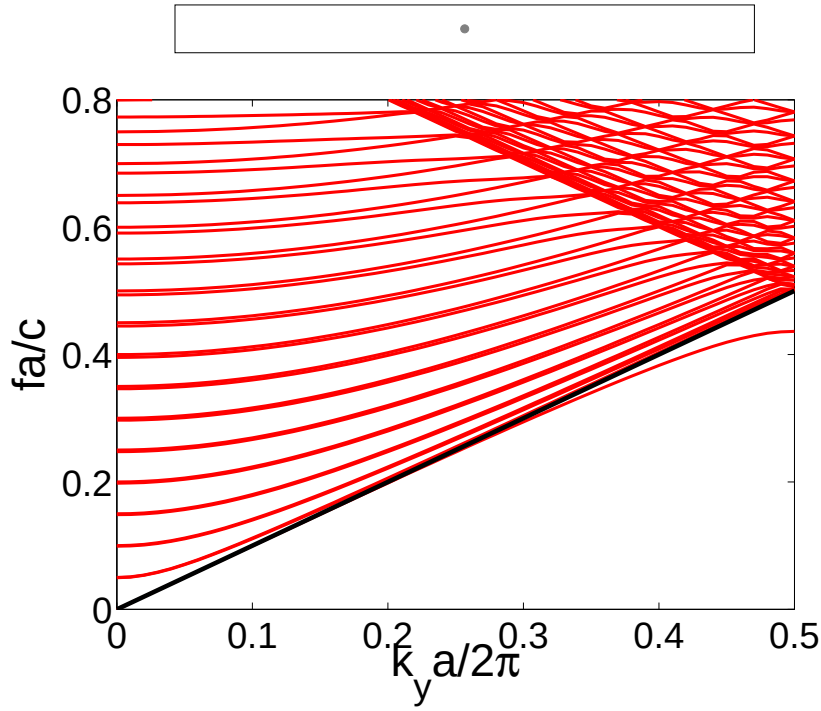

Figure 2.12 Band structure of one layer of dielectric rods. The top figure shows the calculation area with size of 220 mm × 11 mm. The rods have circular cross section with radius r = 1.83 mm and relative permittivity ϵ = 9.8. The bottom shows the band structures based on the supercell method. The black line is the light line.

composed by circular rods with radius r = 1.83 mm and relative permittivity ϵ = 9.8. We can find new modes appearing below the light line in the band gap. They are below the light line, so they are evanescent waves in air; they are in the band gap, so they are evanescent wave in the photonic crystal. So these modes are surface modes. Since the surface modes decay in both the air and the photonic crystal layers, if the two layers are wide enough, the surface modes along different surfaces will not interact with each other. So though we are simulating a system with infinite number of surfaces, the surface modes converges to the surface modes along a single surface. If I change the length of the photonic crystal layer or air layer in the supercell, the locations of other modes may changes in the band structures but the locations of surface modes will stay stable.

2.2.4 Surface modes along a layer of dielectric rods

Using the same supercell method, we can find surface modes exist along a layer of dielectric rods. Fig. 2.12 shows an example. The dielectric rods have circular cross section with radius $r = 1.83$ mm and relative permittivity $\epsilon = 9.8$. The distance between the adjacent rods is $a = 11$ mm. The supercell is a $20a \times a$ rectangle with one rod in the middle. We can find a band under the light line in Fig. 2.12. These modes are evanescent in air, so they are the surface modes.

2.3 Extraordinary transmission and beaming

In previous sections I have demonstrated the localized property of surface modes. However surface modes can also influence the far field since they can propagate at the surface and connect the structures at the surface. Extraordinary transmission (EOT) and beaming are two examples.

2.3.1 Extraordinary transmission

Considering an opaque screen with some open area, the extraordinary transmission means the normalized transmission is larger than 1. The normalized transmission means the transmission which is normalized to the opening area. Suppose the incident intensity is I_{in} with unit W/m^2. If the open area has finite area A, the total transmitted power is P_{out} with unit W. Then the normalized transmission T_{norm} is

$$T_{norm} = \frac{P_{out}}{I_{in}A}.$$

If the open area is a period structure, suppose the area of one unit cell is U and the the area of the open area in one unit cell is A, then the transmission is described by the transmitted intensity I_{out} and the normalized transmission is

$$T_{norm} = \frac{I_{out}U}{I_{in}A}.$$

When the dimension of the holes are much larger than the wavelength, we enter the region of geometrical optics. Based on our everyday experience about shadows, we know $T_{norm} \approx 1$ in this case.

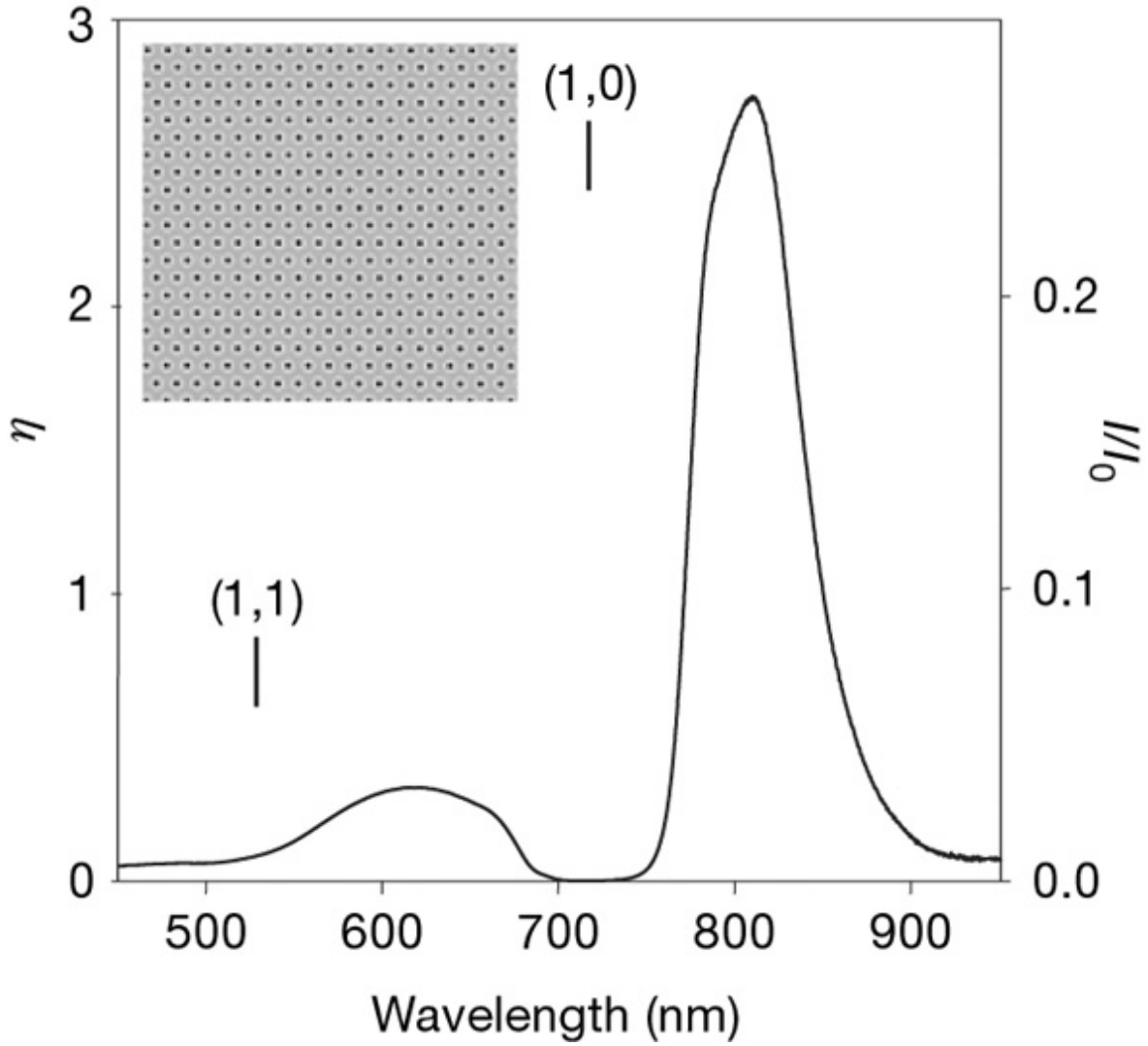

Figure 2.13 Figure excerpted from the paper Nature **445**, 39 (2007) by C. Genet and T. W. Ebbesen. The normalized transmission spectrum through a 225 nm-thick Au film with a triangle lattice of circular holes under normal incidence. The hole diameter is 170 nm and the lattice constant is 520 nm. Before and after the film are the glass substrate and index-matching liquid. η is the normalized transmission.

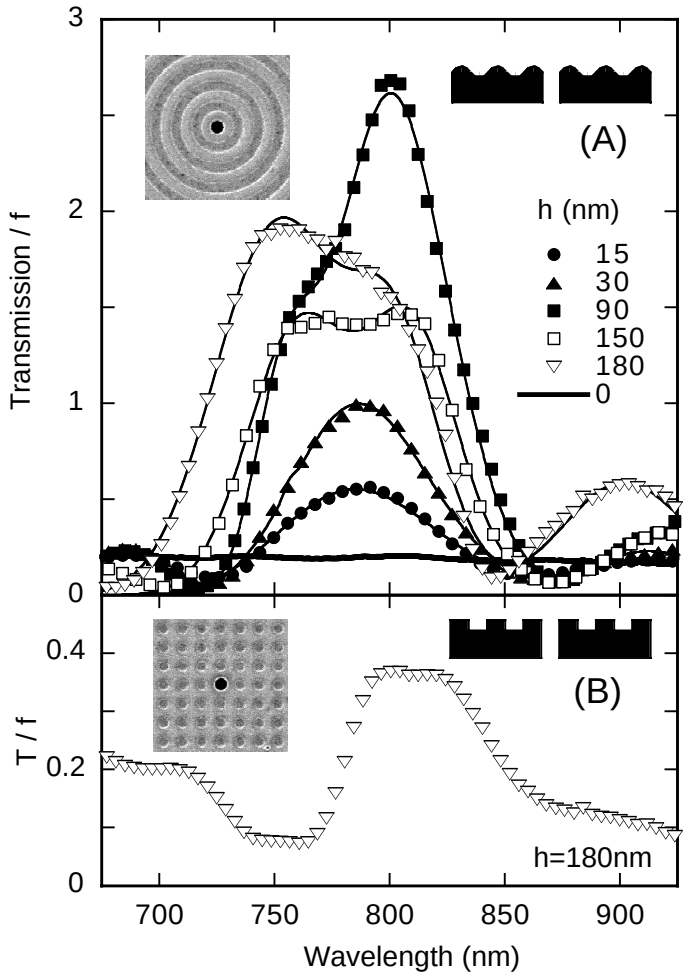

Figure 2.14 Figure excerpted from the paper Opt. Lett. **26**,1972 (2001) by T Thio *et al.*. The normalized transmission spectrum through a Ag hole surrounded by (A) rings and (B) a square lattice of dimples. The rings have sinusoidal cross section. The mean radii is $R_k = kP$, $P = 750\,nm$ and $k = 1, 2, \cdots, 10$. The peak-to-peak amplitude is h. The dimples have lattice constant 750 nm and depth $h = 180$ nm.

In 1944, Bethe (40) has studied the transmission through a subwavelength circular hole with radius r in an infinitely thin perfect electric conductor (PEC) film. He predicted that

$$T_{norm} = \frac{64(kr)^4}{27\pi^2}.$$

So the normalized transmission should be very small when the wavelength is much bigger than the dimension of the holes.

However, it is found the normalized transmission through subwavelength metallic holes can be bigger than 1 recently (41). This phenomenon can not be explain by conventional theories, so it is named as extraordinary transmission. It is generally believed that surface plasmons play important role here to carry energy to the holes along the surface.

Fig. 2.13 shows the normalized transmission spectrum through a 225 nm-thick Au film with a triangle lattice of circular holes (42). EOT happens at some special wavelengths. A common explanation links the peak wavelengths to the momentum-matching condition(43; 44) . The logic is when the momentum-matching condition is satisfied, the surface plasmons are strongly excited and help the transmission. In this thesis, I will analyze this explanation carefully.

EOT through a single hole was also observed (45). As insets of Fig. 2.14 show, the hole is surrounded by rings or dimples, which will excite surface plasmons and enhance the transmission.

2.3.2 Beaming

Grooves at the input surface result in EOT; grooves at the output surface also result in a special phenomenon, beaming: light emerging from a subwavelength aperture is compressed into a narrow beam. (46).

The principle of the beaming can be explained clearly using photonic crystals (47). The top row of Fig. 2.15 shows the field distribution after a photonic crystal waveguide. Since the waveguide is narrow, the exit of the waveguide is similar to a point source and the output fields are diffracted to all angles. A possible way to converging the field is to use surface modes to generate a “big” source. In the middle row of Fig. 2.15 a surface layer is added after the photonic crystal to support the surface mode. We can see clearly surface modes are excited and propagate along the surface layer. So a big source is achieved. But the surface waves are evanescent waves, they won’t influence the far field distribution.

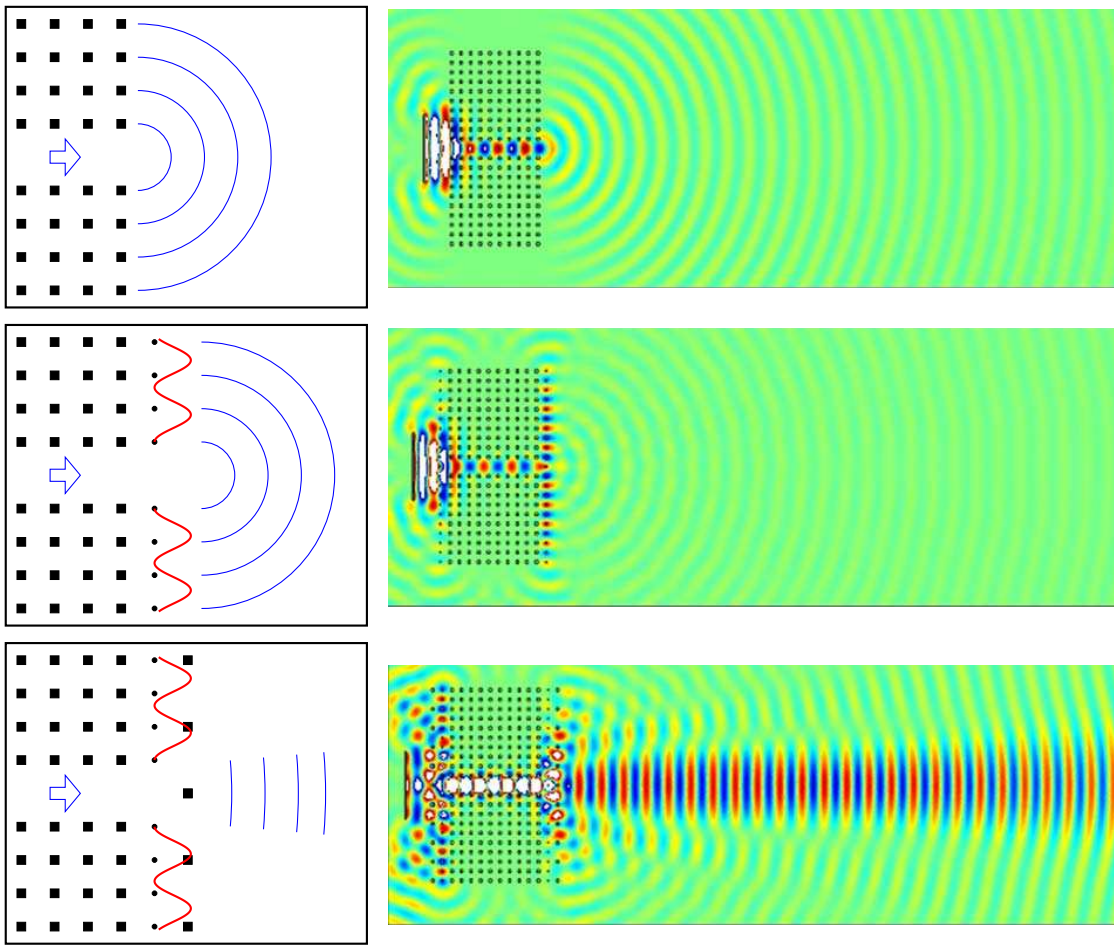

Figure 2.15 Figure excerpted from the paper Phys. Rev. B **76**, 235417 (2007) by R. Moussa *et al.* partially. Left column shows the exit surface of a photonic crystal. From top to bottom: a bare photonic crystal, a photonic crystal with a surface layer, a photonic crystal with a surface layer and grating layer. The right column shows the field distribution after the waveguide.

So another grating layer is necessary to coupling the evanescent wave into propagation wave. As shown in the bottom row of Fig. 2.15, a nice beam is formed after the surface layer and the grating layer. Good beaming can also be achieved by different designs based on the photonic crystals (48; 49; 50; 51; 52).

The metallic beaming structures are also studied experimentally(53; 54) and theoretically(55; 56). They are basically a subwavelength metallic surrounded by periodic structures such as grooves or dielectric gratings along the surface. The flat metal surface supports the surface mode, so additional surface layer is not necessary. The periodic structures works as the grating layer.

CHAPTER 3. Surface Waves Along a Metal-Dielectric Interface with Defects

3.1 Background

In the previous chapter I have proved the existence of surface plasmons at a flat metal-dielectric interface. But the perfect flat metal-dielectric interface rarely exists in reality. A nature question is what the special properties the defects at the surface will bring to the surface waves.

Before the in-depth discussion, I want to clarify two nomenclatures. In this thesis, "surface mode" is a term describing specially the electromagnetic waves bounded at a surface and decaying off exponentially normal to the surface; "surface wave" is a general phrase which just represents the fields along a surface.

Surface plasmons at a flat interface can not be excited by incident plane waves directly because of momentum mismatch (1; 2). Introducing defects at the interface is an efficient method to excite surface plasmons. But the inner mechanics still need be clarified. The existence of the surface plasmons doesn't necessarily mean it will be excited. Zero field is always another existing eigenmode. Even we are convinced the excitation of the surface plasmons, we still want to understand the phase and amplitude of the excited surface plasmons. All these arguments make the analytic deduction form Maxwell's equations valuable.

In 2006, two papers published in Nature Physics tried to explain the same data by two different theories (57; 58). They studied the transmission through a subwavelength slit assisted by a groove. The geometry is shown in the inset of Fig. 3.1. This is 2 two-dimensional system. The inset shows the $x-y$ plane and the system is invariable along z direction. The incident wave is TE polarized.

The experimental, numerical and analytical transmission curves are shown in Fig. 3.1. The experimental curve oscillates with d, the distance between the slit and the groove, which can be explained by the excitation of surface waves. The first order process is: When the incident wave hits the groove,

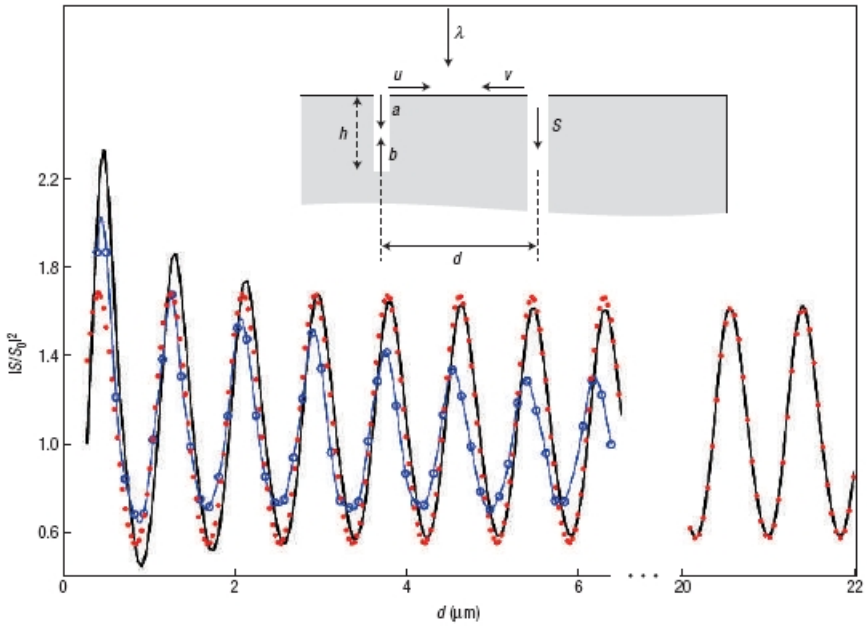

Figure 3.1 Figure excerpted from the paper Nature Phys. **2**, 551 (2006) by P. Lalanne and J. P. Hugonin. The transmission through the subwavelength metallic slit accompanied by a grooves gotten from experiments, numerical simulations and analytic calculations. Blue circles: experimental data. Solid blue line: analytic results based on CDEW model. Solid black line: numerical data computed by the fully vectorial modal method. Red dots: analytic results based on surface plasmon mode. The Inset: the geometry of the metallic slit-groove structure.

the surface wave is excited, propagates at the metal surface and interferes with the field at the slit's entrance. So the transmission is modulated with d. More careful analysis will consider the surfaces wave excited by the slit and the reflected surface wave by the slit and groove. The multiple scattering process will introduce the similar d dependence.

To get good fitting results when d is small, Gay *et al.* (57) gave up the surface plasmon explanation and developed the composite diffracted evanescent waves (CDEW) mode. In the theory, the surface wave excited by an indentation is called CDEW with the form $H_z(y) \propto (\kappa/y + \mu)\cos(k_{surf}y + \phi)$ (the indentation locates at $y = 0$).

In the other paper (58), the authors insisted in the surface plasmon explanation. They simulates the same structure numerically and used the surface plasmon mode to fit the transmission. The fitting is very good when b is big but fails when d is small. Further theory work decomposed (58; 59; 60; 61) the surface wave into two parts: $H_z = H_{sp} + H_c$. H_{sp} is the surface plasmon and H_c is the residual quasicylindrical wave. It was verified numerically that $|H_c| \propto 1/y^{1/2}$ when y is small (58; 59).

In this chapter, I calculate the surface wave along the metal-dielectric interface with indentations from Maxwell's equations analytically. I support the decomposition $H_z = H_{sp} + H_c$ and study the asymptotic behavior of the quasicylindrical wave. I also explain the connection between the surface waves along the metal surface and PEC surface.

3.2 Surface impedance boundary conditions

Since fields decay rapidly inside a metal, we don't care what in the metal and we can use a boundary condition to "replace" it. When the permittivity of the metal is very negative, surface impedance boundary condition can be used to describe the metal-dielectric interface. This boundary condition says the fields on the interface satisfy (20)

$$\vec{E}_{//} = Z_m \hat{n} \times \vec{H}_{//}$$

Here $\hat{n}$ is a unit vector normal to the interface pointing into the dielectric; Z_m is the impedance of the metal and $Z_m = \sqrt{\mu_m \mu_0 / \epsilon_m \epsilon_0}$; μ_m and ϵ_m are the relative permittivity and permeability of the metal.

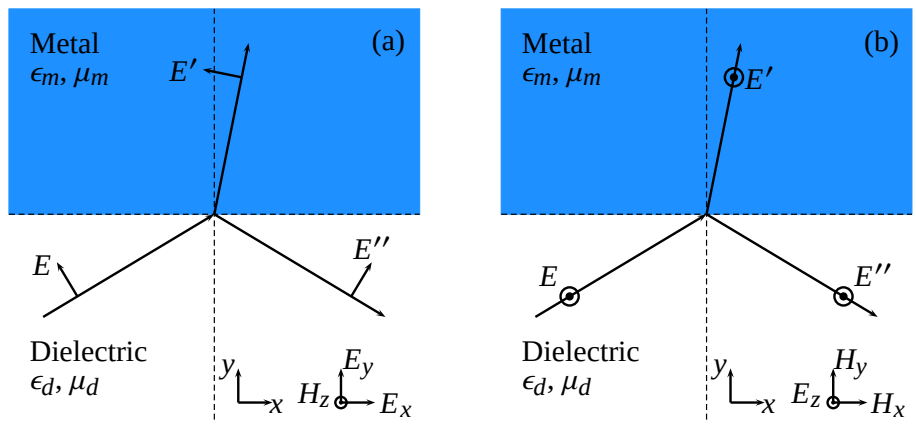

Figure 3.2 Schematic representation of light incidence at a metal-dielectric interface under (a) TE polarization and (b) TM polarization.

3.2.1 Validity of SIBC

In two-dimensional space, we can verify SIBC analytically. Now we consider the plane wave incidence at a metal-dielectric interface and Fig. 3.2 shows the system analyzed. $y = 0$ is the metal-dielectric interface. The relative permittivity and permeability of the dielectric are ϵ_d and μ_d; of the dielectric are ϵ_m and μ_m. The refractive indexes of the dielectric and metal are $n_d = \sqrt{\epsilon_d \mu_d}$ and $n_m = \sqrt{\epsilon_m \mu_m}$. The impedance of the dielectric and metal are $Z_d = \sqrt{\mu_d \mu_0 / \epsilon_d \epsilon_0}$ and $Z_m = \sqrt{\mu_m \mu_0 / \epsilon_m \epsilon_0}$. The incident plane wave is $\vec{E} \exp[i(k \cos\theta x + k \sin\theta y - \omega t)]$ and the reflected wave is $\vec{E}'' \exp[i(k \cos\theta x - k \sin\theta y - \omega t)]$; the corresponding magnetic fields are $\vec{H} \exp[i(k \cos\theta x + k \sin\theta y - \omega t)]$ and $\vec{H}'' \exp[i(k \cos\theta x - k \sin\theta y - \omega t)]$. θ is the incident angle. k is the wavevector in the dielectric and $k = \omega \sqrt{\epsilon_d \epsilon_0 \mu_d \mu_0}$.

We first consider TE polarization as shown in Fig. 3.2(a). Now $\vec{E} = -E \cos\theta \hat{x} + E \sin\theta \hat{y}$, $\vec{E}'' = E'' \cos\theta \hat{x} + E'' \sin\theta \hat{y}$, $\vec{H} = H_z \hat{z}$, $\vec{H}'' = H_z'' \hat{z}$. Here $\hat{n} = -\hat{y}$, so the parallel components are along x and z directions. Consider the boundary conditions along the interface (20), we can get

$$\frac{E''}{E} = \frac{(\mu_d/\mu_m) n_m^2 \cos\theta - n_d \sqrt{n_m^2 - n_d^2 \sin^2\theta}}{(\mu_d/\mu_m) n_m^2 \cos\theta + n_d \sqrt{n_m^2 - n_d^2 \sin^2\theta}}; \tag{3.1}$$

$$E_x = -E \cos\theta; \qquad E_x'' = E'' \cos\theta \tag{3.2}$$

$$H_z = E/Z_d; \qquad H_z'' = E''/Z_d \tag{3.3}$$

If $\mu_m = \mu_d = 1$ and ϵ_m is very negative, which means $|\epsilon_m| \gg |\epsilon_d|$, we have $n_m^2 - n_d^2 \sin^2\theta \approx n_m^2$. At the interface

$$\vec{E}_{//} = -\hat{x}(E - E'')\cos\theta; \qquad \vec{H}_{//} = \hat{z}(E + E'')/Z_d. \tag{3.4}$$

We have $\hat{n} \times \hat{z} = -\hat{x}$ and

$$\begin{aligned}\frac{(E - E'')\cos\theta}{(E_z + E_z'')/Z_d} &= Z_d \cos\theta \frac{n_d\sqrt{n_m^2 - n_d^2\sin^2\theta}}{n_m^2\cos\theta} \\ &\approx \frac{Z_d n_d n_m}{n_m^2} = Z_m.\end{aligned}$$

So we proved SIBC is correct under TE polarization.

When the electric field is perpendicular to the incident plane, we have $\vec{E} = E_z\hat{z}$, $\vec{E}'' = E_z''\hat{z}$, $\vec{H} = H_x\hat{x} + H_y\hat{y}$,$\vec{H}'' = H_x''\hat{x} + H_y''\hat{y}$ and

$$\frac{E_z''}{E_z} = \frac{n_d\cos\theta - (\mu_d/\mu_m)\sqrt{n_m^2 - n_d^2\sin^2\theta}}{n_d\cos\theta + (\mu_d/\mu_m)\sqrt{n_m^2 - n_d^2\sin^2\theta}}; \tag{3.5}$$

$$H_x = E_z\cos\theta/Z_d; \qquad H_x'' = -E_z''\cos\theta/Z_d \tag{3.6}$$

At the interface

$$\vec{E}_{//} = \hat{z}(E_z + E_z''); \qquad \vec{H}_{//} = \hat{x}(E_z - E_z'')\cos\theta/Z_d; \tag{3.7}$$

$$\hat{n} = -\hat{y}. \tag{3.8}$$

We have $\hat{z} = \hat{n} \times \hat{x}$ and

$$\begin{aligned}\frac{E_z + E_z''}{(E_z - E_z'')\cos\theta/Z_d} &= \frac{Z_d}{\cos\theta}\frac{n_d\cos\theta}{(\mu_d/\mu_m)\sqrt{n_m^2 - n_d^2\sin^2\theta}} \\ &\approx \frac{Z_d n_d}{(\mu_d/\mu_m)n_m^2} = Z_m.\end{aligned}$$

So we proved SIBC is also correct under TM polarization. In two-dimensional space, fields are decoupled to TE and TM polarizations. So we verified SIBC is a good approximation for any incident wave in two-dimensional space.

3.2.2 Surface plasmons based on SIBC

We have verified the existence of surface plasmons at a metal-dielectric interface. Now we will try to find surface modes when the metal is replaced by the surface impedance boundary condition.

Considering the two-dimensional $x-y$ plane. Suppose $y = 0$ is the metal-dielectric interface. $y > 0$ is the dielectric area. H_z, E_x, E_y are the three non-zero components of electromagnetic waves for TE polarization. Define the magnetic field in the dielectric as

$$H_z = e^{i(k_x x + k_y y)}. \tag{3.9}$$

Then

$$E_x = -\frac{k_y}{\epsilon_d \epsilon_0 \omega} H_z. \tag{3.10}$$

SIBC says

$$E_x \hat{x} = \sqrt{\frac{\mu_0}{\epsilon_m \epsilon_0}} \hat{y} \times H_z \hat{z}. \tag{3.11}$$

So we get

$$k_y = -\frac{\omega}{c} \frac{\epsilon_d}{\sqrt{\epsilon_m}}; \tag{3.12}$$

$$k_x = \frac{\omega}{c} (\epsilon_d - \frac{\epsilon_d^2}{\epsilon_m})^{1/2}. \tag{3.13}$$

The rigid expression of the dispersion relation is

$$k_x = \frac{\omega}{c} (\frac{\epsilon_d \epsilon_m}{\epsilon_d + \epsilon_m})^{1/2} \tag{3.14}$$

When $|\epsilon_d/\epsilon_m| \ll 1$, we have

$$\frac{\epsilon_d \epsilon_m}{\epsilon_d + \epsilon_m} \approx \epsilon_d - \frac{\epsilon_d^2}{\epsilon_m}. \tag{3.15}$$

So the dispersion relation of SIBC is the first-order approximation of the dispersion relation of real surface plasmons.

In conclusion SIBC also supports surface plasmons.

3.2.3 Waveguide modes based SIBC

Now we begin to study the eigenmode in a two-dimensional waveguide bounded by metal. Fig. 3.3 shows the metallic waveguide we are studying. The width of the waveguide is W; $y = 0$ is the central

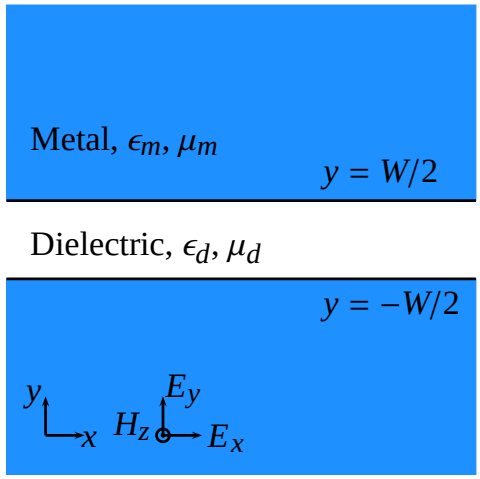

Figure 3.3 Schematic representation of a two-dimensional metallic waveguide.

line of the waveguide; $y = \pm W/2$ are the two metal-dielectric interfaces. $\mu_m = \mu_d = 1$. Only TE mode is studied here.

In the previous chapter, I have gotten the dispersion relations of the modes for the waveguide:

$$\frac{\epsilon_2 k_1}{\epsilon_1 k_2} = \frac{e^{-k_1 W/2} + e^{k_1 W/2}}{e^{-k_1 W/2} - e^{k_1 W/2}} \quad \text{or} \quad \frac{\epsilon_2 k_1}{\epsilon_1 k_2} = \frac{e^{-k_1 W/2} - e^{k_1 W/2}}{e^{-k_1 W/2} + e^{k_1 W/2}}.$$

Then I replace the metal with SIBC. Then the calculation is simpler. Suppose the field in the waveguide is

$$H_z = e^{i\beta x}(Ae^{-k_d y} + Be^{k_d y});$$
$$k_d = \sqrt{\beta^2 - \epsilon_d k_0^2}.$$

Then

$$E_x = -\frac{1}{i\epsilon_d \epsilon_0 \omega}\frac{\partial H_z}{\partial y} = -\frac{ik_d}{\epsilon_d \epsilon_0 \omega} e^{i\beta y}(Ae^{-k_d x} - Be^{k_d x}). \tag{3.16}$$

At the upper interface $y = W/2$, SIBC is $E_x = -Z_m H_z$, so

$$\frac{i\sqrt{\epsilon_m} k_d}{\epsilon_d k_0} = \frac{Ae^{-k_d W/2} + Be^{k_d W/2}}{Ae^{-k_d W/2} - Be^{k_d W/2}}. \tag{3.17}$$

At the lower interface $y = -W/2$, SIBC is $E_x = Z_m H_z$, so

$$\frac{i\sqrt{\epsilon_m} k_d}{\epsilon_d k_0} = -\frac{Ae^{k_d W/2} + Be^{-k_d W/2}}{Ae^{k_d W/2} - Be^{-k_d W/2}}. \tag{3.18}$$

Repeat the calculation about the metal-dielectric-metal multilayer structure and we get the dispersion relation

$$\frac{i\sqrt{\epsilon_m}k_d}{\epsilon_d k_0} = \frac{e^{-k_d W/2} + e^{k_d W/2}}{e^{-k_d W/2} - e^{k_d W/2}} \quad \text{or} \quad \frac{i\sqrt{\epsilon_m}k_d}{\epsilon_d k_0} = \frac{e^{-k_d W/2} - e^{k_d W/2}}{e^{-k_d W/2} + e^{k_d W/2}}. \tag{3.19}$$

When $\beta \gg \sqrt{\epsilon_d}k_0$, these modes are high-order evanescent modes of the waveguide which decay very quickly. We focus the basic mode which satisfies $\beta \lessapprox \sqrt{\epsilon_d}k_0$. If $|\epsilon_m| \gg |\epsilon_d|$, we have

$$k_m = \sqrt{\beta^2 - \epsilon_m k_0^2} \approx -i\sqrt{\epsilon_m}k_0 \Longrightarrow \frac{\epsilon_m k_d}{\epsilon_d k_m} \approx \frac{i\sqrt{\epsilon_m}k_d}{\epsilon_d k_0}. \tag{3.20}$$

I choose $\sqrt{-1} = -i$ here to ensure $\mathrm{Re}(k_m) > 0$. So the dispersion relation of the basic mode in a SIBC waveguide converges to real metal waveguide when ϵ_m is very negative.

In the next chapter I will study the transmission through metallic waveguides. The knowledge about these eigenmodes will be very useful.

3.2.4 Summary of SIBC

As the calculation above shows, SIBC is a good approximation to describe metal-dielectric interfaces when $|\epsilon_m\mu_m| \gg |\epsilon_d\mu_d|$. It makes analytic deduction simpler and saves the calculating area of numeric simulations. When $|\epsilon_m|$ goes to infinity, the fields calculated from SIBC converge to the real fields.

3.3 Green's function

3.3.1 Surface wave at a metal-dielectric interface with defects

Fig. 3.4 shows the structure we are studying: a dielectric-metal interface with grooves and slits. The blue area is metal and the white area is dielectric. $x = 0$ is the metal-dielectric interface with some subwavelength indentations. The relative permittivity and permeability of the dielectric are ϵ_d and μ_d; The relative permittivity and permeability of the metal are ϵ_m and μ_m. Normally $\mu_d = \mu_m = 1$. Define ϵ_0 and μ_0 as the permittivity and permeability of vacuum. Then the impedance of the dielectric and metal are $Z_d = \sqrt{\mu_d\mu_0/\epsilon_d\epsilon_0}$ and $Z_m = \sqrt{\mu_m\mu_0/\epsilon_m\epsilon_0}$.

Surface impedance boundary condition is used to describe the metal-dielectric interface. At the metal-dielectric interface $x = 0$ in Fig. 3.4, the boundary condition is expressed as $E_y + Z_m H_z = 0$ when

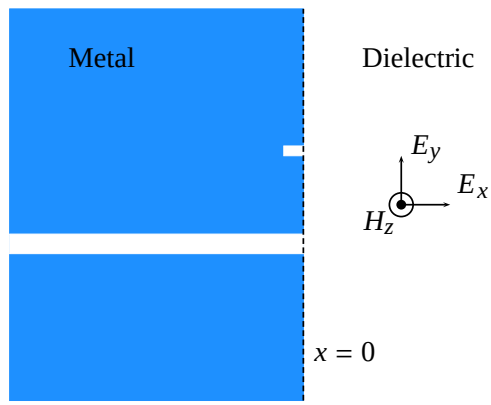

Figure 3.4 Schematic representation of a dielectric-metal interface with indentations in a two-dimensional space.

the fields are TE polarized. I have shown SIBC is precise when $|\epsilon_m \mu_m| \gg |\epsilon_d \mu_d|$. Silver has relative permittivity $\epsilon_m = -8.5 + 0.76i$ at $\lambda = 500$ nm (λ is the wavelength in free space) (62). So the silver-air interface can be described by SIBC up to 500 nm. When λ becomes smaller, $|\epsilon_m|$ becomes smaller and we have to give up the approximation. When λ increases, $|\epsilon_m|$ increases rapidly and the fields calculated from SIBC converge to the real fields.

Suppose no incident waves coming from the dielectric side in Fig. 3.4, using the mode expansion method (63), we can prove the surface fields along $x = 0$ satisfy

$$H_z(y) = \frac{1}{Z_d \lambda_d} \int_{-\infty}^{\infty} dy' \Big(E_y(y') + Z_m H_z(y') \Big) G(y, y'). \tag{3.21}$$

The Green's function is

$$G(y, y') = G(y - y') = \int_{-\infty}^{\infty} dk \frac{e^{ik(y-y')}}{k_x + Z_s k_d}. \tag{3.22}$$

Here $Z_s = Z_m / Z_d$ and $Z_d = \sqrt{\mu_d \mu_0 / \epsilon_d \epsilon_0}$, $\arg(Z_s) \in (-\pi/2, -\pi/4)$. $Z_s = 0$ when the metal is PEC. k_d and λ_d is the wave vector and the wavelength in the dielectric. k_x is the x component of the wave vector of the plane waves in the $x > 0$ half-plane. The plane waves propagate or decay along $+x$ direction, so

$$k_x = \begin{cases} \sqrt{k_d^2 - k^2} & \text{if } k_d \geq k \\ i\sqrt{k^2 - k_d^2} & \text{if } k_d < k \end{cases} \tag{3.23}$$

Since $E_y(y') + Z_m H_z(y') = 0$ at the metal surface, Eq. (3.21) is equivalent to the expression

$$H_z(y) = \frac{1}{Z_d \lambda_d} \int_{Indentation} dy' \Big(E_y(y') + Z_m H_z(y') \Big) G(y, y'). \tag{3.24}$$

So the surface wave is decided by the fields in the indentations. When the width of an indentation is much smaller than the wavelength, the fields at the exit of the indentation are close to a delta function. Then the Green's function describes the surface wave excited by the indentation. If the indentation is wide, we have to know the fields inside the indentation to calculate the surface wave. Though the surface wave around the indentation is complex, the Green's function still describes the surface wave in the region more than several widths away from the indentation.

Above we have shown the surface wave along the output surface of slits. It is easy to prove the fields along the input surface have the similar form. They are the summation of the surface wave excited by indentations plus some terms related to the incident and reflected waves. The surface wave excited by indentations can also be calculated by other methods (64), but the Green's function method uses less assumptions and gets precise results. We have no need to assume the existence of surface plasmons beforehand. They will emerge from the Green's function automatically.

3.3.2 Simplification of the Green's function

Define $y = s\lambda_d, k = hk_d$, then

$$G(y) = g(y/\lambda) = g(s) = \int_{-\infty}^{\infty} dh \frac{e^{i2\pi hs}}{\sqrt{1-h^2} + Z_s}.$$

Now let's focus on the non-dimensional Green's function $g(s)$. The square root function is defined as $\arg(\sqrt{1-h^2}) \in (-\pi/2, \pi/2]$ because of Eq. (3.23). $g(s)$ is an even function. We choose $s > 0$ here.

When $s \to 0$, $g(s)$ diverges but

$$\int_{-\infty}^{\infty} dh \Big(\frac{e^{i2\pi hs}}{\sqrt{1-h^2} + Z_s} - \frac{e^{i2\pi hs}}{\sqrt{1-h^2}} \Big)$$

does not. So

$$g(s) \to \pi H^{(1)}_{2\pi s} \text{ when } s \to 0. \tag{3.25}$$

We moves to the complex space of h to do the integral in $g(s)$. The integration path is shown in Fig. 3.5 as the red curves. The blue lines are the branch cut to make $\sqrt{1-h^2}$ single-valued. The function $\sqrt{1-h^2}$ is not continuous along the the path $\mathrm{Im}(h) = 0$ at the point $h = 1$ under this branch cut. So the integration path has two loops.

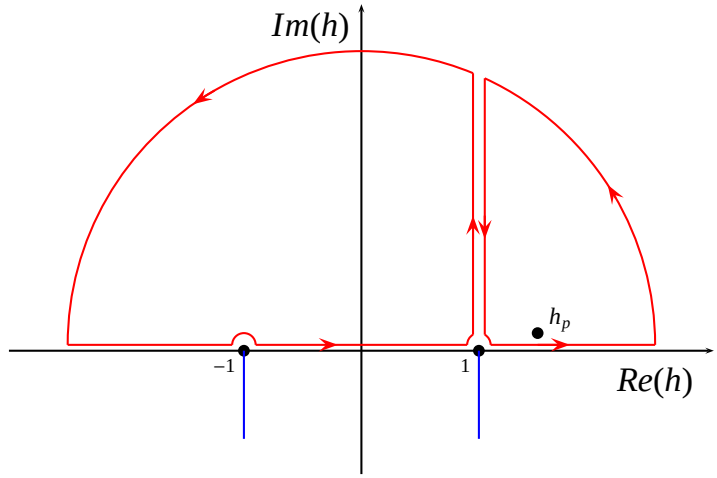

Figure 3.5 Integration path of the non-dimensional Green's function in the complex space of h. The red curves are the integration path and the blue lines are the branch cut.

Using Cauchy's integral formula to the left and right loops, we get

$$\int_{-\infty}^{1} dh \frac{e^{i2\pi hs}}{\sqrt{1-h^2}+Z_s} + \int_0^{\infty} dq \frac{ie^{i2\pi s}e^{-2\pi qs}}{\sqrt{q^2-2qi}+Z_s} = 0;$$
$$\int_{\infty}^{0} dq \frac{ie^{i2\pi s}e^{-2\pi qs}}{-\sqrt{q^2-2qi}+Z_s} + \int_1^{\infty} dh \frac{e^{i2\pi hs}}{\sqrt{1-h^2}+Z_s} = 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s}$$

Here h_p is a pole: $\sqrt{1-h_p^2}+Z_s = 0$. The integral variable is changed from h to q along the vertical parts of the path, $h = 1+qi$ and $q \in [0,\infty)$. Again $\arg(\sqrt{1-h^2}) \in (-\pi/2, \pi/2]$, $\arg(\sqrt{q^2-2qi}) \in (-\pi/2, \pi/2]$.

Then

$$g(s) = 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s} + 2ie^{i2\pi s} \int_0^{\infty} dq e^{-2\pi qs} \frac{\sqrt{q^2-2qi}}{Z_s^2 - q^2 + 2qi}.$$

In the second term, the integrand contributes significantly only when $q \in (0, 1/s)$. So we neglect q^2 when $s \gtrsim 1$. Then

$$g(s) \approx 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s} + 2ie^{i2\pi s} \int_0^{\infty} dq e^{-2\pi qs} \frac{\sqrt{-2qi}}{Z_s^2 + 2qi}. \tag{3.26}$$

All the calculations below base on the approximation expression of Eq. (3.26), which is good when $s > 1$. But our numeric simulation results show the approximation works very well even when s is smaller than 1. We will also prove that Eq. (3.26) gives the same asymptotic result as Eq. (3.25) when s is very small. So the approximation in Eq. (3.26) is precise for almost any s.

The first term in Eq. (3.26) is the well-understood surface plasmon $g_{SP}(s)$. The second is the residual quasicylindrical wave $g_C(s)$ (61), which need further study. Define

$$I(s) = \int_0^\infty dq e^{-2\pi q s} \frac{\sqrt{q/Y}}{Y+q}; \qquad Y = Z_s^2/(2i). \tag{3.27}$$

Then

$$I(s) = \sqrt{\frac{1}{2Ys}} - \pi e^{2\pi Y s} \mathrm{Erfc}(\sqrt{2\pi Y s}); \tag{3.28}$$

$$g_C(s) = \sqrt{Y}(1-i)e^{i2\pi s} I(s). \tag{3.29}$$

Here $\arg(\sqrt{Y}) \in (\pi/4, \pi/2)$; Erfc is the complementary error function; $|I(s)|$ describes the envelope of the quasicylindrical wave.

The integrand of $I(s)$ in Eq. (3.27) contributes significantly only when $0 < q \lesssim 1/s$. When $|Y|s \gg 1$, q in the interval $(0, 1/s)$ satisfies

$$\begin{aligned} \frac{1}{Y+q} &\approx \frac{1}{Y}(1 - \frac{q}{Y}) \\ \Rightarrow I(s) &\approx \frac{1}{4\sqrt{2}\pi} \frac{1}{(Ys)^{3/2}} - \frac{1}{16\sqrt{2}\pi^2} \frac{1}{(Ys)^{5/2}}. \end{aligned}$$

When $|Y|s$ is small, $I(s)$ can be simplified by the Taylor expansion of $e^{u^2}\mathrm{Erfc}(u)$,

$$e^{u^2}\mathrm{Erfc}(u) = 1 - \frac{2}{\sqrt{\pi}} u + u^2 + \cdots$$

We can call $|Y|s$ as the surface distance because it describes how the envelope of the quasicylindrical wave evolves during the propagation. The piecewise function below gives good estimation about $I(s)$.

$$I(s) \approx \begin{cases} \sqrt{1/2Ys} & \text{if } |Y|s < 0.002 \\ \sqrt{1/2Ys} - \pi & \text{if } |Y|s < 0.02 \\ 1/[4\sqrt{2}\pi(Ys)^{3/2}] & \text{if } |Y|s > 1 \end{cases}$$

When $0.02 < |Y|s < 1$, we have to include more terms of the Taylor expansion or we can calculate Erfc function directly.

We now compare the amplitude of g_{SP} and g_C. Since g_{SP} decays exponentially, g_C will be stronger for sure when s is very big. But in practice, h_p is close to 1 with a very small imaginary part for good

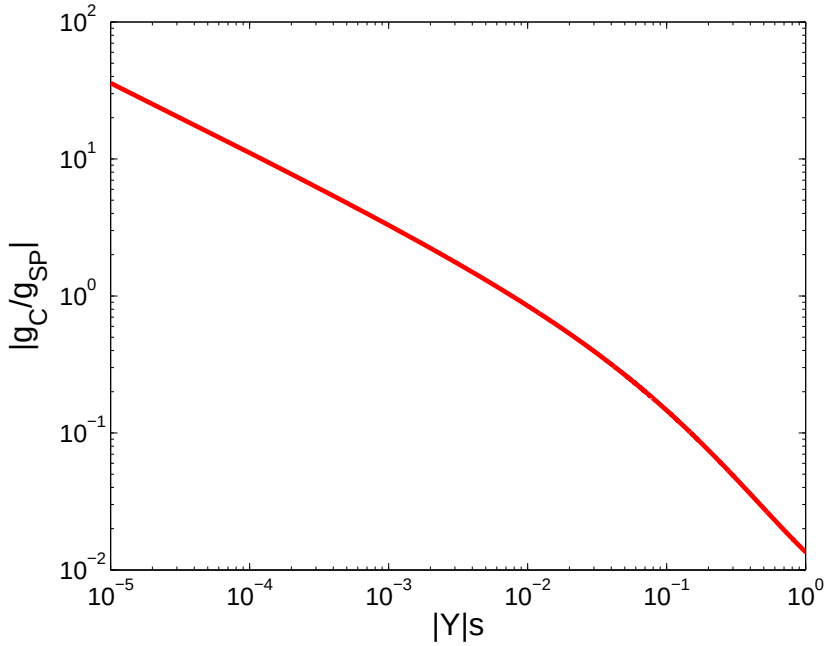

Figure 3.6 $|g_C(s)/g_{SP}(s)|$ along a metal-air interface when the metal permittivity is $\epsilon_m = -31.39 + 2.22i$.

noble metals. So $|g_{SP}(s)/g_C(s)| \approx |I(s)/2\pi|$. Fig. 3.6 shows $|g_{SP}(s)/g_C(s)|$ along a metal-air interface when $\epsilon_m = -31.39 + 2.22i$, which is the relative permittivity of silver at $\lambda = 852$ nm (62). If the metal is lossless, $I(s)$ is a function of $|Y|s$, which is the horizontal axis in Fig. 3.6. We can find the quasicylindrical wave is negligible when $|Y|s > 0.1$.

3.3.3 Discussion

Put the results together, we have

$$g_{SP}(s) = 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s};$$

$$g_{C0}(s) = (1-i)e^{i2\pi s}\frac{1}{\sqrt{2s}}$$

$$g_{C1}(s) = -i\pi Z_s e^{i2\pi s}$$

$$g_{CB}(s) = Z_s \frac{i}{4\sqrt{2}\pi}\frac{e^{i2\pi s}}{(Ys)^{3/2}}$$

The piecewise function below gives good approximation about the Green's function.

$$g(s) \approx \begin{cases} g_{SP}(s) + g_{C0}(s) & \text{if } |Y|s < 0.002 \\ g_{SP}(s) + g_{C0}(s) + g_{C1}(s) & \text{if } |Y|s < 0.02 \\ g_{SP}(s) + g_{CB}(s) & \text{if } |Y|s > 0.1 \end{cases}$$

The metal-dielectric interface is divided into three regions based on the surface distance: the near region ($|Y|s < 0.02$), the intermediate region ($0.02 < |Y|s < 0.1$) and the far region ($|Y|s > 0.1$).

If the metal is nearly lossless, g_{SP} dominates in the far region. If the metal is a PEC, $Z_s = 0$ and $Y = 0$. The whole surface belongs to the near region and

$$g(s) \approx g_{C0} = (1 - i)e^{i2\pi s}\frac{1}{\sqrt{2s}}.$$

It agrees with the PEC's Green's function with Hankel function form (65). When s is very small, $|g_{C0}| \gg |g_{SP}|$, we can neglect g_{SP} and return to the result in Eq. (3.25).

The fields along a metal-air interface are calculated numerically using the commercial finite element method software COMSOL Multiphysics. The simulation area is similar to Fig. 5.9 but without the groove. The incident wave comes through the slit in $+x$ direction. The fields along the output surface are recorded and compared with analytic results in Fig. 3.7 and 3.8. The slit width is 0.05λ. The relative error of the simulations is around 2%. The analytic results in the figures are calculated by Eq. (3.24) by replacing $G(y, y')$ with the different simplified forms of the Green's function shown in the legends. We assume that $E_y(y') + ZH_z(y')$ is constant at the exit of the slit and the value is taken from the numerical simulations. There is no fitting parameters in the analytic calculations.

In Fig. 3.7 and 3.8, the real and imaginary parts of H_z are plotted in the range $y \in (0.1\lambda, 40\lambda)$ at two incident wavelengths. We always plot the simulation results along with the analytic results from $g_{SP} + g_C$ and the two curves always overlap. It shows the approximation in Eq. (3.26) is good in the region $y > 0.1\lambda$.

Fig. 3.7 shows the fields when $\lambda = 3000$ nm and $\epsilon_m = -329 + 47.5i$. The near region is the area $y \lesssim 12\lambda$. The analytic results from $g_{SP} + g_{C0} + g_{C1}$ agree with the simulation results very well in this region and g_{C1} component do improve the analytic results. When s enters the intermediate region, results from $g_{SP} + g_{C0}$ and $g_{SP} + g_{C0} + g_{C1}$ have visible difference from the simulation results. The

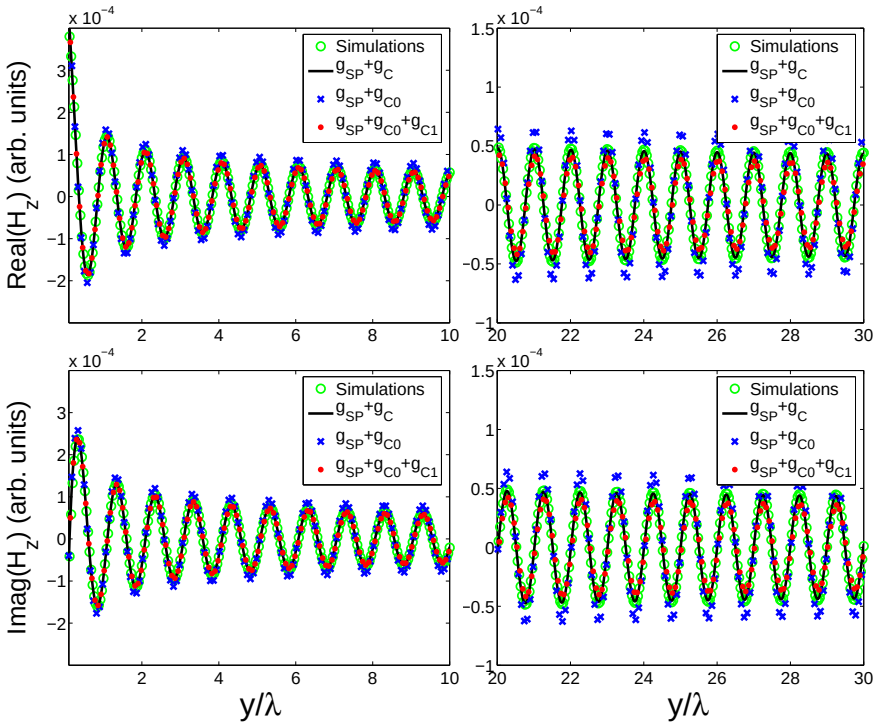

Figure 3.7 Real and imaginary parts of H_z along a silver-air interface in the regions $y \in (0.1\lambda, 10\lambda)$ and $y \in (20\lambda, 30\lambda)$ when $\lambda = 3000$ nm and $\epsilon_m = -329 + 47.5i$. The simulation results are calculated by COMSOL Multiphysics, a commercial finite element method software. The other curves are calculated analytically using different simplified forms of the Green's function shown in the legends.

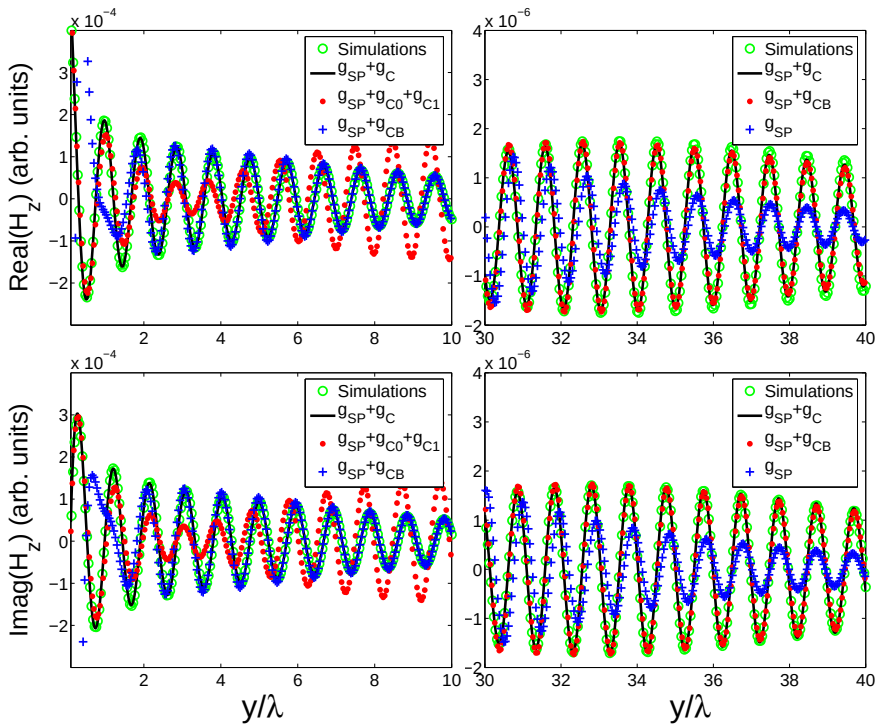

Figure 3.8 Real and imaginary parts of H_z along a metal-air interface in the regions $y \in (0.1\lambda, 10\lambda)$ and $y \in (30\lambda, 40\lambda)$ when $\lambda = 500$ nm and $\epsilon_m = -8.50 + 6.00i$. The simulation results are calculated by COMSOL Multiphysics. The other curves are calculated analytically using different simplified forms of the Green's function shown in the legends.

difference looks small here because g_{SP} is strong. In Fig. 3.8, the incident wavelength is $\lambda = 500$ nm. The permittivity of the metal is set as $\epsilon_m = -8.50 + 6.00i$ with a big imaginary part to suppress the the surface plasmon in the far region. Since $|Y|$ is big here, the near and intermediate regions are short. $g_{SP} + g_{CB}$ gives good analytic results when $y > 2\lambda$. Fig. 3.8 shows clearly that the quasicylindrical wave decays as $y^{-3/2}$ in the far region.

Our analytic calculation also proves $g_{SP}(s) + g_{C0}(s)$ is a very good approximation of $g(s)$ for a good noble metal at visible wavelengths (59; 60; 61). It is the approximation we used in the near region. In the far region, we have proved $g_C(s)$ is negligible, so is $g_{C0}(s)$. In the intermediate region, this approximation is poor. But the intermediate region is short and it is difficult to notice the fitting error.

In EOT research, the interesting surface area is normally between $\lambda/2$ and 100λ. For a good noble metal at visible wavelengths, the area belongs to the far region and surface plasmon dominates. When wavelength increases and ϵ_m is more negative, the metal surface converges to the PEC surface in two ways:(i)the near region becomes longer and (ii)g_{C0} becomes stronger comparing with g_{SP} in the near region. The whole surface works like a mixture of PEC and metal surfaces and $|Y|$ servers as a good index to describe the mixture state. The reference (58) shows the converging process graphically. Both metal and PEC surfaces support the strong slow-decaying surface waves, so they have similar phenomena such as EOT.

3.3.4 Conclusion

In this section, we have analyzed the surface wave along metal-dielectric interface with an indentation. We get the asymptotic forms of the wave far away from and close to the indentation. Based on the surface distance $|Y|s$, the interface is divided into several regions and the surface wave behaves differently in every region. The complete description of the surface wave would give us deeper understanding about light transmission mechanics through metallic apertures.

CHAPTER 4. Theory of Extraordinary Transmission

4.1 Background

In the previous chapters I have introduce EOT through a 2D periodic subwavelength hole arrays or a single subwavelength hole surrounded by structures. Served as a simpler model, EOT through 1D slits have also been demonstrated for both periodic slits (66; 67; 68) and one slits surrounded by grooves (69). The common explanation linking the transmission peaks with the momentum-matching condition is also used to explain the transmission peaks here (66). But another paper verified numerically that when the momentum-matching condition is fulfilled, the transmission has dips actually (67).

In this section I will present an theory about the transmission through 1D subwavelength periodic metallic slits by first principles. I will prove analytically that transmission dips appear when momentum-matching condition is satisfied. I will also study the relations between transmission peaks and geometry parameters.

4.2 Mode expansion method

Mode expansion method is used to calculate the transmission. In this section I study the plane wave incidence to infinitely-long slits and use it as a example to demonstrate the method. Fig. 4.1 shows the schematic picture we will study. The whole system is a two-dimensional structure. The $x < 0$ half-plane is free space and the $x > 0$ half-plane is metal with several infinitely-long slits. For simplicity the medium in the slits is air too. To excite the surface plasmons along the metal surface, the incident plane wave is TE polarized.

The fields in the free space include the incident plane wave and reflected waves. The incident wave

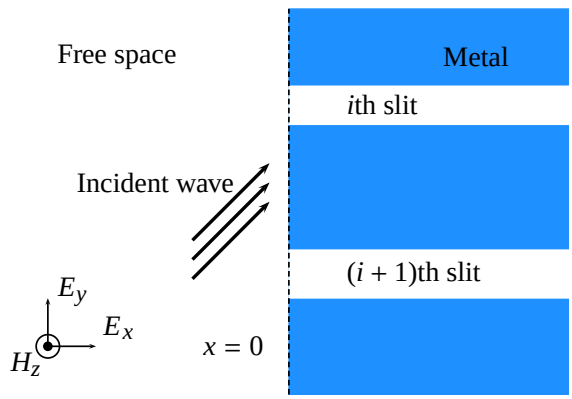

Figure 4.1 Schematic picture to study the transmission into infinitely-long slits.

is

$$H_z^{Inc}(x, y) = \exp[i(k_\perp x + k_{//} y)];$$

$$k_\perp = k_0 \cos\theta; \qquad k_{//} = k_0 \sin\theta.$$

Here k_0 is the wavevector of the incident wave in free space and θ is the incident angle.

The reflected wave is expressed as the superposition of all possible plane waves.

$$H_z^{Ref} = \int_{-\infty}^{\infty} dk \rho(k) \exp[i(-k_x x + ky)];$$

$$k_x = \begin{cases} \sqrt{k_0^2 - k^2} & \text{if } k_0 \geq k \\ i\sqrt{k^2 - k_0^2} & \text{if } k_0 < k \end{cases}$$

The reflected plane waves either propagates along $-x$ direction or decay along $-x$ direction, so k_x is either positive real number or pure imaginary number with positive imaginary part.

So the fields in the free space are

$$H_z(x, y) = \exp[i(k_\perp x + k_{//} y)] + \int_{-\infty}^{\infty} dk \rho(k) \exp[i(-k_x x + ky)];$$

$$E_y(x, y) = \frac{1}{\epsilon_0 \omega}\left(k_\perp \exp[i(k_\perp x + k_{//} y)] - \int_{-\infty}^{\infty} dk k_x \rho(k) \exp[i(-k_x x + ky)]\right).$$

The field in the slits can be expressed as the summation of eigenmodes of the slits. Suppose $\Phi_{ij}(x, y)$ is the magnetic field of the jth eigenmodes in the ith slit. Then

$$\Phi_{ij}(x, y) = \phi_{ij}(y) \exp[i\beta_{ij} x].$$

The eigenmodes in a metallic slit have been calculated in the previous chapter. Normally

$$\phi_{ij}(y) = A\exp[i\alpha_{ij}(y-y_i)] + B\exp[-i\alpha_{ij}(y-y_i)] \text{ and } \alpha_{ij}^2 + \beta_{ij}^2 = \frac{\omega^2}{\epsilon_0^2\mu_0^2}.$$

Here A and B are two complex coefficients; $y = y_i$ is the central line of the slit. If the metal is lossless, α_{ij} and β_{ij} are real numbers or pure imaginary numbers. Since the slits are infinitely long, there is no output surface and obviously no reflection wave in the slits. So the eigenmodes either propagate in $+x$ direction or decays in $+x$ direction. So β_{ij} is either a positive real number or a pure imaginary number with positive imaginary part.

So the fields in the slits are

$$H_z(x,y) = \sum_i \sum_j A_{ij}\phi_{ij}(y)\exp[i\beta_{ij}x];$$
$$E_y(x,y) = \frac{1}{\epsilon_0\omega}\sum_i \sum_j A_{ij}\beta_{ij}\phi_{ij}(y)\exp[i\beta_{ij}x].$$

Next step I will use the boundary conditions along $x = 0$ to calculate the unknown coefficients $\rho(k)$ and A_{ij}.

Using the field distribution in free space, $E_y - ZH_z$ at $(0, y)$ is

$$E_y(0,y) - Z_m H_z(0,y) = (\frac{k_\perp}{\epsilon_0\omega} - Z_m)\exp[ik_{//}y] - \int_{-\infty}^{\infty} dk(\frac{k_x}{\epsilon_0\omega} + Z_m)\rho(k)\exp[iky].$$

Using the field distribution in the slit, we can get $E_y - ZH_z$ at the entrance of a slit is

$$E_y(0,y) - Z_m H_z(0,y) = \sum_i \sum_j A_{ij}(\frac{\beta_{ij}}{\epsilon_0\omega} - Z_m)\phi_{ij}(y).$$

H_z and E_y are continuous at the entrances of the slits, so is $E_y - ZH_z$. So when $(0, y)$ is at the entrance of a slit, we have

$$(\frac{k_\perp}{\epsilon_0\omega} - Z_m)\exp[ik_{//}y] - \int_{-\infty}^{\infty} dk(\frac{k_x}{\epsilon_0\omega} + Z_m)\rho(k)\exp[iky] = \sum_i \sum_j A_{ij}(\frac{\beta_{ij}}{\epsilon_0\omega} - Z_m)\phi_{ij}(y).$$

Surface impedance boundary condition requires $E_y - Z_m H_z = 0$ at the metal surface. So when $(0, y)$ is at the metal surface,

$$(\frac{k_\perp}{\epsilon_0\omega} - Z_m)\exp[ik_{//}y] - \int_{-\infty}^{\infty} dk(\frac{k_x}{\epsilon_0\omega} + Z_m)\rho(k)\exp[iky] = 0.$$

We can define naturally $\phi_{ij}(y) = 0$ when the point $(0, y)$ is not in the ith slit. So the equation

$$(\frac{k_\perp}{\epsilon_0\omega} - Z_m)\exp[ik_{/\!/}y] - \int_{-\infty}^{\infty} dk(\frac{k_x}{\epsilon_0\omega} + Z_m)\rho(k)\exp[iky] = \sum_i\sum_j A_{ij}(\frac{\beta_{ij}}{\epsilon_0\omega} - Z_m)\phi_{ij}(y).$$

is correct for any y.

Using inverse Fourier transformation, we will get the expression of $\rho(k)$.

$$(\frac{k_x}{\epsilon_0\omega} + Z_m)\rho(k) = (\frac{k_\perp}{\epsilon_0\omega} - Z_m)\delta(k - k_{/\!/}) - \frac{1}{2\pi}\sum_i\sum_j A_{ij}(\frac{\beta_{ij}}{\epsilon_0\omega} - Z_m)\int_{-\infty}^{\infty} dy\phi_{ij}(y)\exp[-iky] \quad (4.1)$$

H_z is continuous at the entrance of the slits, so when $(0, y)$ at the entrance of a slit,

$$\exp[ik_{/\!/}y] + \int_{-\infty}^{\infty} dk\rho(k)\exp[iky] = \sum_i\sum_j A_{ij}\phi_{ij}(y). \quad (4.2)$$

Now considering an eigenmode in the i'th slit $\Phi_{i'j'} = \phi_{i'j'}(y)\exp[i\beta_{i'j'}x]$. Suppose the upper and lower boundaries of the i'th slit is $y = w_1$ and $y = w_2$. Obviously when $y \in (w_2, w_1)$, Eq. (4.2) is correct. So

$$\int_{w_1}^{w_2} \phi^*_{i'j'}(y)\Big(\exp[ik_{/\!/}y] + \int_{-\infty}^{\infty} dk\rho(k)\exp[iky]\Big)dy = \int_{w_1}^{w_2} \phi^*_{i'j'}(y)\Big(\sum_i\sum_j A_{ij}\phi_{ij}(y)\Big)dy.$$

Since we have defined $\phi_{i'j'}(y) = 0$ when $y \notin (w_2, w_1)$, we can extend the integral interval to $(-\infty, \infty)$.

$$\int_{-\infty}^{\infty} \phi^*_{i'j'}(y)\exp[ik_{/\!/}y]dy + \int_{-\infty}^{\infty} dy\int_{-\infty}^{\infty} dk\rho(k)\phi^*_{i'j'}(y)\exp[iky] = \sum_i\sum_j A_{ij}\int_{-\infty}^{\infty} \phi^*_{i'j'}(y)\phi_{ij}(y)dy. \quad (4.3)$$

The eigenmodes are orthogonal with each other,

$$\int_{-\infty}^{\infty} \phi^*_{i'j'}(y)\phi_{ij}(y)dy = 0 \qquad \text{if } i \neq i' \text{ or } j \neq j'.$$

When $i \neq i'$, the two slits don't overlap, so $\phi^*_{i'j'}(y) = 0$ when $\phi^*_{ij}(y)$ is not zero and vice versa. When $i = i'$, it is because of the orthogonality of eigenmodes of a waveguide. So the right term of Eq. (4.3) equals

$$\sum_i\sum_j A_{ij}\int_{-\infty}^{\infty} \phi^*_{i'j'}(y)\phi_{ij}(y)dy = A_{i'j'}P_{i'j'};$$

and here $P_{i'j'}$ is the norm square of $\phi_{i'j'}(y)$

$$\int_{-\infty}^{\infty} \phi^*_{i'j'}(y)\phi_{i'j'}(y)dy = P_{i'j'};$$

Now let's focus on the second term in Eq. (4.3). Using Eq. (4.1) to replace $\rho(k)$ in this term and adjusting the order of the integration, we will get

$$\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dk \rho(k) \phi^*_{i'j'}(y) \exp[iky] = \frac{k_\perp - Z_m \epsilon_0 \omega}{k_\perp + Z_m \epsilon_0 \omega} \int_{-\infty}^{\infty} \phi^*_{i'j'}(y) \exp[ik_{//} y] dy$$
$$- \sum_i \sum_j A_{ij} (\beta_{ij} - Z_m \epsilon_0 \omega) \frac{1}{2\pi} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dy' \phi^*_{i'j'}(y) \phi_{ij}(y) \int_{-\infty}^{\infty} dk \frac{\exp[ik(y-y')]}{k_x + Z_m \epsilon_0 \omega}$$

It is easy to verify $Z_m \epsilon_0 \omega = Z_s k_0$.

Finally for the jth eigenmode in the ith slit, Eq. (4.3) becomes

$$F_0 I_{ij} - \sum_{i'} \sum_{j'} A_{i'j'} (\beta_{i'j'} - Z_s k_0) G_{ij,i'j'} = A_{ij} P_{ij}. \tag{4.4}$$

Here I exchange the notations of i, j and i', j' and

$$F_0 = \frac{2k_\perp}{k_\perp + Z_s k_0};$$
$$I_{ij} = \int_{-\infty}^{\infty} \phi^*_{ij}(y) \exp[ik_{//} y] dy;$$
$$P_{ij} = \int_{-\infty}^{\infty} \phi^*_{ij}(y) \phi_{ij}(y) dy;$$
$$G_{ij,i'j'} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dy' \phi^*_{i'j'}(y') \phi_{ij}(y) G(y, y');$$
$$G(y, y') = G(y - y') = \int_{-\infty}^{\infty} dk \frac{\exp[ik(y-y')]}{k_x + Z_s k_0}.$$

Eq. (4.4) is the key equations from the mode expansion method, which is a self-consistent linear equations about A_{ij}. There are infinite eigenmodes in a slit. Theoretically we have to include all the eigenmodes and the summation of j' goes from 0 to ∞. In practice we can only include finite eigenmodes in Eq. (4.4). For every eigenmode we consider, we can write a corresponding equation. Then we can solve the coefficients of such eigenmodes by Eq. (4.4) and get the field in the slits. The field in the free space can be calculated by Eq. (4.1) furthermore.

For the equation corresponding to the eigenmode ϕ_{ij} in Eq. (4.4), the first term is the influence of the incident wave; the term on the right describe the property of the eigenmode itself. The second term is the most interesting one. I have shown the Green's function $G(y - y')$ describes the surface wave excited by an indentation in the previous chapter. So the second term describes the interaction between two eigenmodes through the surface waves. This term confirms the key role of surface waves in the transmission explicitly.

The dynamic picture of surface waves is like this: the incident wave comes to the a defect at the metal-dielectric interface and a surface wave is excited; then the surface wave propagates along the the surface and evolves continuously; then the surface wave hits another defect and change the field distribution in the second defect; some surface wave is scattered, some is reflected and the rest goes through the defect and continue to propagate forwards; then the reflected and transmitted surface wave will meet other defects and this process repeats. Actually Eq. (4.4) describes the multiple scattering process of surface waves. The surface waves propagates forward and backward between defects. Finally they will reach a stable state and Eq. (4.4) is the conditions that the equilibrium state satisfies.

4.3 Transmission into a 1D subwavelength grating with infinite length

4.3.1 Simplification of the general equations

In the previous section we have gotten the self-consistent linear equations about the field inside the slits. These equations are quite general, regardless the widths and locations of the slits. If there are m slits and we consider n eigenmodes in each slit. We have $m \times n$ unknown coefficients and Eq. (4.4) provides $m \times n$ equations.

In this section I will study the transmission through a one-dimensional subwavelength grating illuminated by a normal incident plane wave. Suppose the period of the grating is a and the width of the slits is W. We can simplify Eq. (4.4) dramatically in this special situation.

For a general case with m slits and n eigenmodes in each slit, there are $m \times n$ equations in Eq. (4.4). A grating has an infinite number of slits. But they are all equivalent with each other under normal incidence. The fields in one slit equal the field in all the other slits. So we only need write down the equations corresponding the eigenmodes in one slit, which are n equations. We choose to study the fields in the zeroth slit here.

The incident angle is $\theta = 0°$, so $k_\perp = k_0$ and $k_{/\!/} = 0$. Then

$$F_0 = \frac{2}{1+Z_s};$$
$$I_{0j} = \int_{-\infty}^{\infty} \phi_{0j}^*(y)dy;$$
$$P_{0j} = \int_{-\infty}^{\infty} \phi_{0j}^*(y)\phi_{0j}(y)dy;$$
$$G_{0j,i'j'} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dy' \phi_{0j'}^*((y' - a(i'-1))\phi_{0j}(y)G(y,y').$$

I neglect the first subscript of these variables which represents the index of the slits, and define $A_j = A_{0j}, \beta_j = \beta_{0j}, I_j = I_{0j}$ and $P_j = P_{ij}$, then for the jth eigenmode in the slit, Eq. (4.4) becomes

$$F_0 I_j - \sum_{i'=-\infty}^{\infty}\sum_{j'=0}^{n-1} A_{j'}(\beta_{j'} - Z_s k_0)G_{0j,i'j'} = A_j P_j. \tag{4.5}$$

In Eq. (4.5), i' goes from $-\infty$ to ∞ since a grating has an infinite number of slits. I consider the first n eigenmodes in a slit, so j' goes from 0 to $n-1$ and there are n equations in Eq. (4.5). Here I am studying the subwavelength slits. The slit is very narrow and there is only one propagation mode in the slit. I will only consider the propagation mode here. $n = 1$ and j' changes from 0 to 0. Define

$$G_{i'} = G_{00,i'0} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dy' \phi_{00}^*((y' - i'a)\phi_{00}(y)G(y,y').$$

Then Eq. (4.5) becomes

$$F_0 I_0 - A_0(\beta_0 - Z_s k_0)\sum_{i'=-\infty}^{\infty} G_{i'} = A_0 P_0. \tag{4.6}$$

Eq. (4.6) includes only one equation and one unknown coefficient A_0, which can be solved directly.

$$A_0 = \frac{F_0 I_0}{P_0 - (\beta_0 - Z_s k_0)\sum_{n=-\infty}^{\infty} G_n} \tag{4.7}$$

Here i' is a dummy variable, so I change it to n for simplicity.

The first eigenmode of a slit is $\Phi_0(x,y) = \phi_0(y)\exp[i\beta_0 x]$. When the slit is bounded by PEC, $\beta_0 = k_0$ and $\phi_0(y)$ is a constant; I set $\phi_0(y) = 1$ here. For a SIBC slit, if the slit is very narrow comparing with the wavelength, $\phi_0(y)$ equals a constant approximately. If the metal is lossless, $\beta_0 > k_0$; otherwise it is a complex number. In summary, suppose the central line of the zeroth slit is $y = 0$, we can assume the basic eigenmode of a very narrow slit has

$$\phi_0(y) = \begin{cases} 1 & \text{when} |y| \le W/2; \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$
\begin{aligned}
I_0 &= \int_{-W/2}^{W/2} dy = W; \\
P_0 &= \int_{-W/2}^{W/2} dy = W; \\
G_n &= \frac{1}{2\pi} \int_{-W/2}^{W/2} dy \int_{-W/2}^{W/2} dy' G(y - y' + an); \\
G(y) &= \int_{-\infty}^{\infty} dk \frac{\exp[ik(y - y')]}{k_x + Z_s k_0}.
\end{aligned}
$$

We have analyzed $G(y)$ detailed in the previous section. $G(y)$ is an even function, so $G_n = G_{-n}$. $G(y)$ diverges when y goes to 0. So G_0 is an integration around infinity. When y is away from 0, $G(y)$ is a smooth function and changes slowly. The variable of $G(y)$ function changes from $na - W$ to $na + W$ in the integration range of G_n. If $n \neq 0$, $W \ll \lambda$, we can assume $G(y) \approx G(na)$ in this range. So

$$
G_n \approx \frac{W^2}{2\pi} G(an) \qquad \text{when } n \neq 0.
$$

and

$$
\sum_{n=-\infty}^{\infty} G_i = G_0 + 2 \sum_{n=1}^{\infty} G_n \approx G_0 + \frac{W^2}{\pi} \sum_{n=1}^{\infty} G(na).
$$

4.3.2 The PEC grating

In the previous section, I studied the properties of the non-dimensional Green's function. Define $s = y/\lambda$, the non-dimensional Green's function

$$
g(s) = G(s\lambda) = G(y).
$$

When the grating is made by PEC, we have $Z_s = 0$ and the Green's function is

$$
g(s) = \int_{-\infty}^{\infty} dk \frac{\exp[iky]}{\sqrt{k_0^2 - k^2}} = \pi H_0^{(1)}(2\pi s) \approx (1 - i) e^{i2\pi s} \frac{1}{\sqrt{2s}}.
$$

So

$$
\sum_{n=1}^{\infty} G(na) \approx (1 - i) \sum_{n=1}^{\infty} e^{i2\pi na/\lambda} \frac{1}{\sqrt{2na/\lambda}}
$$

When $a = \lambda$,

$$\sum_{n=1}^{\infty} G(n\lambda) \approx (1-i)\sum_{n=1}^{\infty}\frac{1}{\sqrt{2n}};$$

$$\text{We know } \sum_{n=1}^{\infty}\frac{1}{\sqrt{2n}} = \infty, \text{ so } \left|\sum_{n=1}^{\infty} G(n\lambda)\right| = \infty.$$

The coefficient of the basic eigenmode is

$$A_0 = \frac{F_0 I_0}{P_0 - \beta_0[G_0 + W^2/\pi \sum_{n=1}^{\infty} G(na)]} \tag{4.8}$$

Obviously, when $a = \lambda$, the denominator goes to infinity and $A_0 = 0$. So no field in the slit and the transmission is 0.

Fig. 4.2 shows the normalized transmission into a SIBC subwavelength grating with infinite length illuminated by a normal incident plane wave. The incident wavelength is $\lambda = 560$ nm and the slit width of the grating is $W = 30$ nm. The transmission for different grating periods a is calculated. The analytic results are calculated using Eq. (4.8) and the simulation results are obtained by COMSOL Multiphysics simulations. We can find clearly when a/λ is an integer, the transmission has a dip. For most of periods, the normalized transmission is bigger than 1, which means EOT.

4.3.3 The SIBC grating

When the grating is made by real metal, I use SIBC to replace the metal. The surface wave along the metal surface has two parts: the surface plasmon and the residue quasicylindrical wave. The Green's function is

$$g(s) = \int_{-\infty}^{\infty} dk \frac{\exp[iky]}{\sqrt{k_0^2 - k^2} + Z_s k_0} \approx g_{SP}(s) + g_C(s).$$

$g_{SP}(s)$ and $g_C(s)$ represent the surface plasmon and the residue quasicylindrical wave.

$$g_{SP}(s) = 2\pi i \frac{Z_s}{h_p} e^{i2\pi h_p s}$$

$$g_C(s) = \sqrt{Y}(1-i)e^{i2\pi s}\Big(\sqrt{\frac{1}{2Ys}} - \pi e^{2\pi Y s}\mathrm{Erfc}(\sqrt{2\pi Y s})\Big)$$

Here h_p is a pole satisfying $\sqrt{1-h_p^2} + Z_s = 0$, $Y = Z_s^2/(2i)$, Erfc is the complementary error function. When s is big,

$$g_C(s) \approx g_{CB}(s) = Z_s \frac{i}{4\sqrt{2\pi}} \frac{e^{i2\pi s}}{(Ys)^{3/2}}.$$

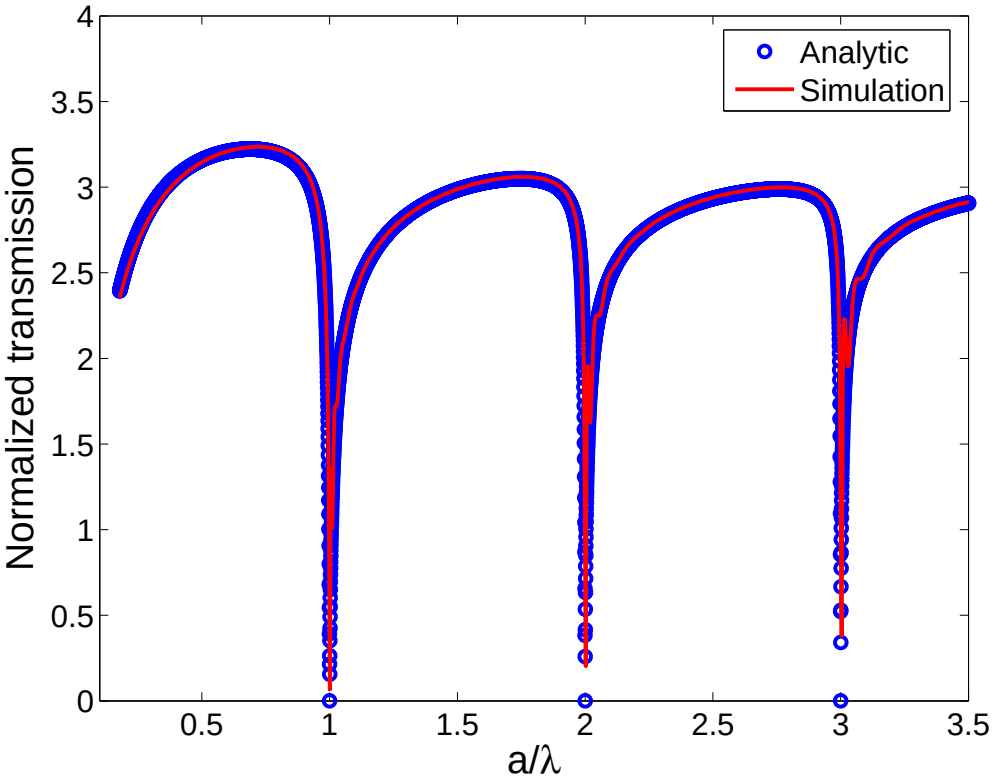

Figure 4.2 The normalized transmission into a PEC subwavelength grating with infinite length illuminated by a normal incident plane wave. The incident wavelength is $\lambda = 560$ nm and the slit width of the grating is $W = 30$ nm.

If the metal is lossless, h_p is a real number bigger than 1. Define $\lambda_{SP} = \lambda/h_p$ as the wavelength of the surface plasmon. When $a = \lambda_{SP}$,

$$\sum_{n=1}^{\infty} G(na) = \sum_{n=1}^{\infty} G(n\lambda_{SP}) \approx 2\pi i \frac{Z_s}{h_p} \sum_{n=1}^{\infty} 1 + \sum_{n=1}^{\infty} g_C(ns_{sp}).$$

The first term on the right side of the approximation diverges obviously. To study the convergence property of the second term, we can use the asymptotic expression of $g_C(s)$. Actually

$$\sum_{n=1}^{\infty} \frac{e^{i2\pi n s_{sp}}}{(nYs_{sp})^{3/2}}$$

always converges. So summation on the right of the approximation diverges. Then the transmission is zero.

Fig. 4.3 shows the normalized transmission into a SIBC subwavelength grating with infinite length illuminated by a normal incident plane wave. The geometry is the same as the PEC grating in Fig. 4.2; the slit width is 30 nm and the grating period a changes. The simulation results in the upper and lower figures of Fig. 4.3 are the same. The analytic curve in the upper figure of Fig. 4.3 is calculated used the precise approximation $g(s) \approx g_{SP}(s) + g_C(s)$, which agrees with the simulation curve very well. In the lower figure, I neglect the residue quasicylindrical wave and assume the surface wave equals the surface plasmon. When a is big, the distance between two neighboring slit is big; $g_C(s)$ becomes small comparing with the surface plasmon in the far region; so the simulation and analytic curves agree with each. When a is small, the distance between two neighboring slits is in near region; $g_C(s)$ becomes big and the analytic results without $g_C(s)$ have big error. All the curves have dips and the curves agree with each other very well. The dips are because the summation of the surface plasmon terms diverges. So we can neglect $g_C(s)$ safely around the dips. At the dips $a = n\lambda_{SP}$; n is an integer.

Combining the results of the PEC and SIBC gratings, I can define the resonant wavelength of the input surface of a metal grating as

$$\lambda_R = \frac{a}{n}\sqrt{1 - Z_s^2}; \qquad n \text{ is an integer.}$$

Then the basic conclusion here is at the resonant wavelengths, the transmission is 0. Actually the the resonant wavelengths are the wavelength that fulfill the momentum-matching conditions (42).

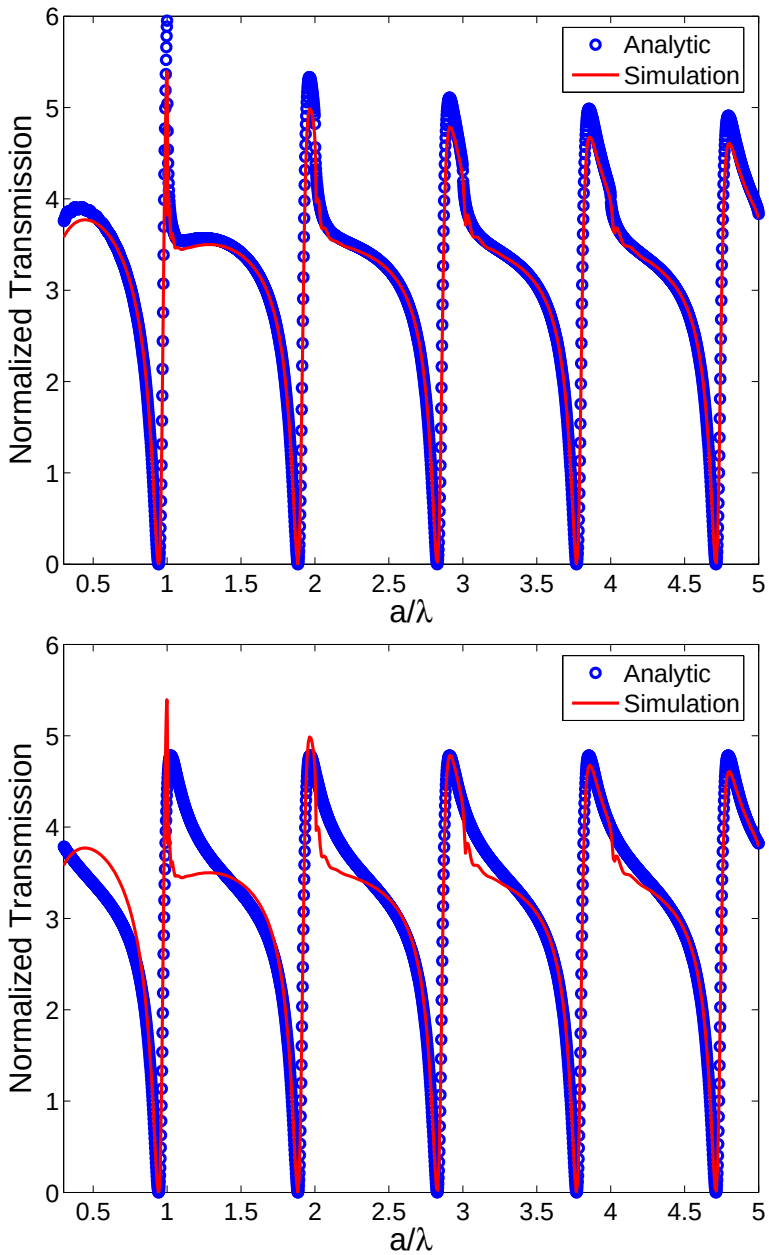

Figure 4.3 The normalized transmission into a SIBC subwavelength grating with infinite length illuminated by a normal incident plane wave. The incident wavelength is $\lambda = 560$ nm, the relative permittivity of the metal is $\epsilon_m = -7.92$, and the slit width of the grating is $W = 30$ nm. The upper and lower figures have the same simulation curve. In the upper curve, the analytic curve is calculated using the approximation of the Green's function $g(s) \approx g_{SP}(s) + g_C(s)$. The analytic curve in the lower figure is calculated using the approximation $g(s) \approx g_{SP}(s)$.

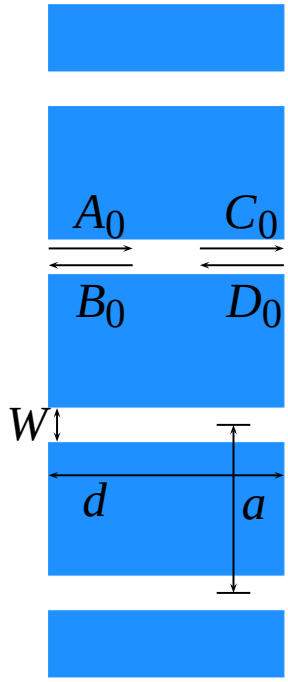

Figure 4.4 Schematic picture to of a metallic grating with a finite length.

4.4 Transmission through a 1D subwavelength grating with finite length

In reality a grating always has finite length. I will study these gratings in the section. I will study not only the transmission dips but also peaks.

The schematic picture of the grating is shown in Fig. 4.4. a and W are the grating period and the slit width, the same as in the previous sections. The slits are subwavelength, so $W \ll \lambda$. d is the length of the grating.

4.4.1 The input surface

I study the input surface first. I also use the mode expansion method. The fields before the grating in free space are the same as the grating with an infinite length . The fields in the slits are different. We have to consider the wave reflected by the output surface. We can assume the magnetic field in the slits as

$$H_z(x, y) = A_0\phi_0(y)\exp[i\beta_0 x] + B_0\phi_0(y)\exp[-i\beta_0 x]. \tag{4.9}$$

This expression includes the forward- and backwards-propagating basic eigenmode of the slit. A_0 and B_0 are unknown coefficients.

Considering the fields H_z and $E_y+Z_mH_z$ along the input surface and repeating the analytic deduction in the previous sections, we can get a similar equation about A_0 and B_0.

$$F_0I_0 - \Big[(A_0 - B_0)\beta_0 - (A_0 + B_0)Z_sk_0)\Big]\sum_n G_n = (A_0 + B_0)P_0. \tag{4.10}$$

At the resonant wavelengths of the input surface, $\sum_n G_n$ diverges.

$$\Big[(A_0 - B_0)\beta_0 - (A_0 + B_0)Z_sk_0)\Big] = 0$$

is the only choice for Eq. (4.10). So at the resonant wavelengths, A_0 and B_0 satisfy

$$\frac{B_0}{A_0} = \frac{\beta_0 - Z_sk_0}{\beta_0 + Z_sk_0}. \tag{4.11}$$

If the metal is lossless, Z_s is a pure imaginary number and β_0 is a positive real number. So $|B_0/A_0| = 1$. It means the energy moving forward equals the energy moving backward. There is no net transmission in the slits. So we get the same conclusion as the grating with an infinite length: the transmission curve has dips at the resonant wavelengths.

I want to extend my analytic calculation to two-dimensional hole arrays in the future. A big difference between a one-dimensional grating and a two-dimensional hole array is that there exists at least one propagation mode in a 1D slit but probably no propagation mode exists in a 2D hole. I will calculate the net power flux carried by evanescent waves here.

Suppose H_z is the only non-zero component of the magnetic field in a hole and

$$H_z(x, y, z) = A_0\phi_0(y, z)\exp[i\beta_0x] + B_0\phi_0(y, z)\exp[-i\beta_0x].$$

β_0 is a pure imaginary number here. Then E_y is

$$E_y(x, y, z) = \frac{\beta_0}{\epsilon_0\omega}\Big(A_0\phi_0(y, z)\exp[i\beta_0x] - B_0\phi_0(y, z)\exp[-i\beta_0x]\Big).$$

The x component of the time-average Poynting vector is

$$\begin{aligned}
< S_x > &= \frac{1}{2}\mathrm{Re}(E_y \times H_z^*) \\
&= \frac{1}{2\epsilon_0\omega}|\phi_0(y, z)|^2(|A_0|^2 - |B_0)|^2)\mathrm{Re}[\beta_0]+ \\
&\quad \frac{1}{2\epsilon_0\omega}|\phi_0(y, z)|^2\exp[2i\beta_0x]\mathrm{Re}[\beta_0A_0B_0^*] + \frac{1}{2\epsilon_0\omega}|\phi_0(y, z)|^2\exp[-2i\beta_0x]\mathrm{Re}[\beta_0A_0^*B_0].
\end{aligned}$$

These terms $1/2\epsilon_0\omega$, $|\phi_0(y,z)|^2$, $(|A_0|^2 - |B_0)|^2)$, $\exp[2i\beta_0 x]$ and $\exp[-2i\beta_0 x]$ are real, so I take them out of the Re function. We also have $\mathrm{Re}[\beta_0] = 0$. So

$$< S_x > = \quad \frac{1}{2\epsilon_0\omega}|\phi_0(y,z)|^2 \exp[2i\beta_0 x]\mathrm{Re}[\beta_0 A_0 B_0^*] + \frac{1}{2\epsilon_0\omega}|\phi_0(y,z)|^2 \exp[-2i\beta_0 x]\mathrm{Re}[\beta_0 A_0^* B_0].$$

At the resonant wavelengths, A_0 and B_0 satisfy Eq. (4.11). Now β_0 and Z_s are pure imaginary numbers and k_0 is a real number. Therefore B_0/A_0 is a real number, so are $A_0 B_0^*$ and $A_0^* B_0$. Then we have $\mathrm{Re}[\beta_0 A_0 B_0^*] = 0$, $\mathrm{Re}[\beta_0 A_0^* B_0] = 0$, and finally $< S_x >= 0$. I proved here that the transmission curve has dips at the resonant wavelength even there are only evanescent modes existing in the holes of a 2D hole arrays.

4.4.2 The output surface

Now I begin to study the output surface. The magnetic filed in the slit is

$$H_z(x,y) = C_0\phi_0(y)\exp[i\beta_0 x] + D_0\phi_0(y)\exp[-i\beta_0 x]$$

and the field in the free space after the grating is

$$H_z(x,y) = \int_{-\infty}^{\infty} dk\rho(k)\exp[i(k_x x + ky)].$$

I move the origin to the output surface and consider the boundary conditions along the surface, then I get the similar equation again

$$\Big[(C_0 - D_0)\beta_0 + (C_0 + D_0)Z_s k_0\Big]\sum_n G_{0n} = (C_0 + D_0)P_0. \tag{4.12}$$

At the resonant wavelengths of the output surface, C_0 and D_0 satisfy

$$\frac{D_0}{C_0} = \frac{\beta_0 + Z_s k_0}{\beta_0 - Z_s k_0}.$$

So the transmission is 0 at the resonant wavelengths again.

In this section I assume the medium before and after the grating is air, so the input surface and output surface have the same resonant wavelengths. It is possible the two surfaces have different resonant wavelengths. For example there is a subtract before the grating. I proved here the transmission is zero at the resonant wavelengths of both the input and the output surfaces.

4.4.3 Transmission through a finite-length grating

I have gotten the equations about the coefficients of the eigenmodes in the slit at the input and output surface and discussed the transmission at the resonant wavelengths. Now I will put them together and calculate the transmission at any wavelengths.

Define

$$R = \frac{(\beta_0 + Z_s k_0)\sum_n G_{0n} - P_0}{(\beta_0 - Z_s k_0)\sum_n G_{0n} + P_0} \tag{4.13}$$

$$T = \frac{F_0 I_0}{(\beta_0 - Z_s k_0)\sum_n G_{0n} + P_0} \tag{4.14}$$

Then

$$A_0 - RB_0 = T; \tag{4.15}$$

$$C_0 R = D_0; \tag{4.16}$$

$$C_0 = A_0 \exp[i\beta_0 d]; \tag{4.17}$$

$$D_0 = B_0 \exp[-i\beta_0 d]. \tag{4.18}$$

Eq. (4.15) is equivalent to Eq. (4.10) and Eq. (4.16) is equivalent to Eq. (4.12). Eq. (4.17) connects A_0 and C_0 since I move the origin when I deduce Eq. (4.12), so does Eq. (4.18).

Solve these equations and I get the transmission through the finite-length grating is

$$\text{Transmission} = \frac{|T|^2}{1-|R|^2}\frac{\beta_0 W}{k_0 a}\left|\frac{1-|R|^2}{1-e^{2id\beta_0}R^2}\right|^2. \tag{4.19}$$

T and R are functions of Z_s, a and W, which are the surface parameters. The only term contain the grating length d is the term at the end of the expression between the absolute value function, which is a typical term of the Fabry-Pérot oscillation term. Its value oscillates with d between 1 and a positive number smaller than 1. Fixing the surface parameters and the incident wavelength, the maximum transmission a grating can reach by changing the grating length is

$$\text{Maximum transmission} = \frac{|T|^2}{1-|R|^2}\frac{\beta_0 W}{k_0 a}. \tag{4.20}$$

From now on, the maximum transmission is defined as the maximum transmission which can be reached by scanning the grating length and fixing all other parameters. Fig. 4.5 shows the maximum

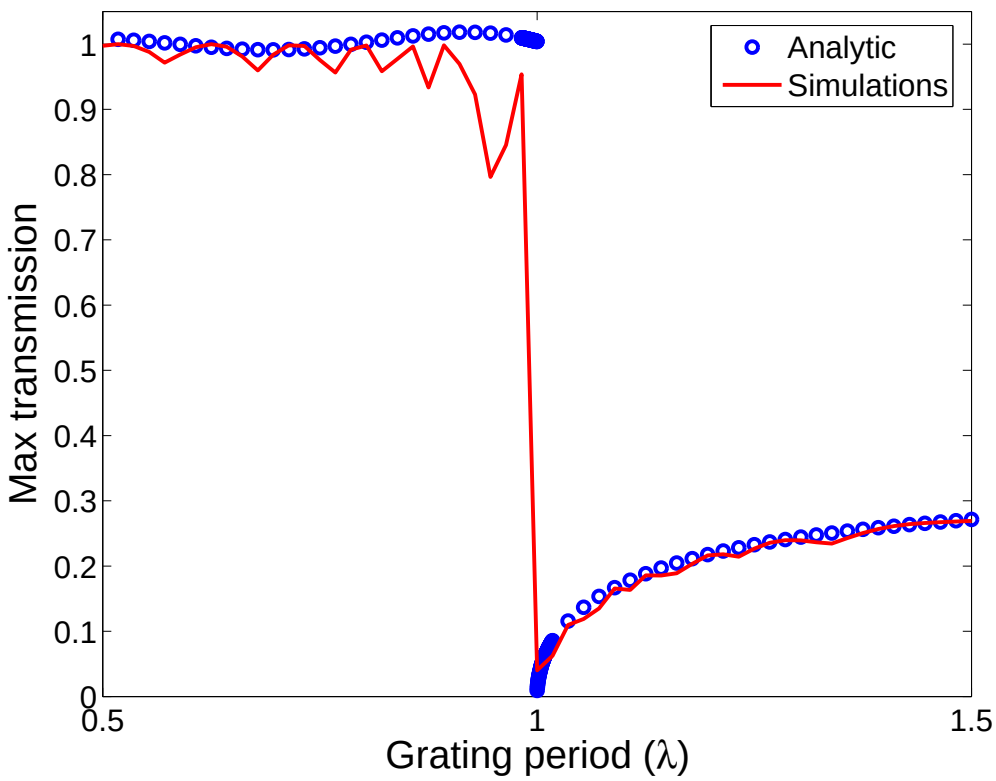

Figure 4.5 The maximum transmission a PEC grating can reach by changing the grating length at different grating periods, The slit width is $W = 30$ nm and the incident wavelength is $\lambda = 560$ nm.

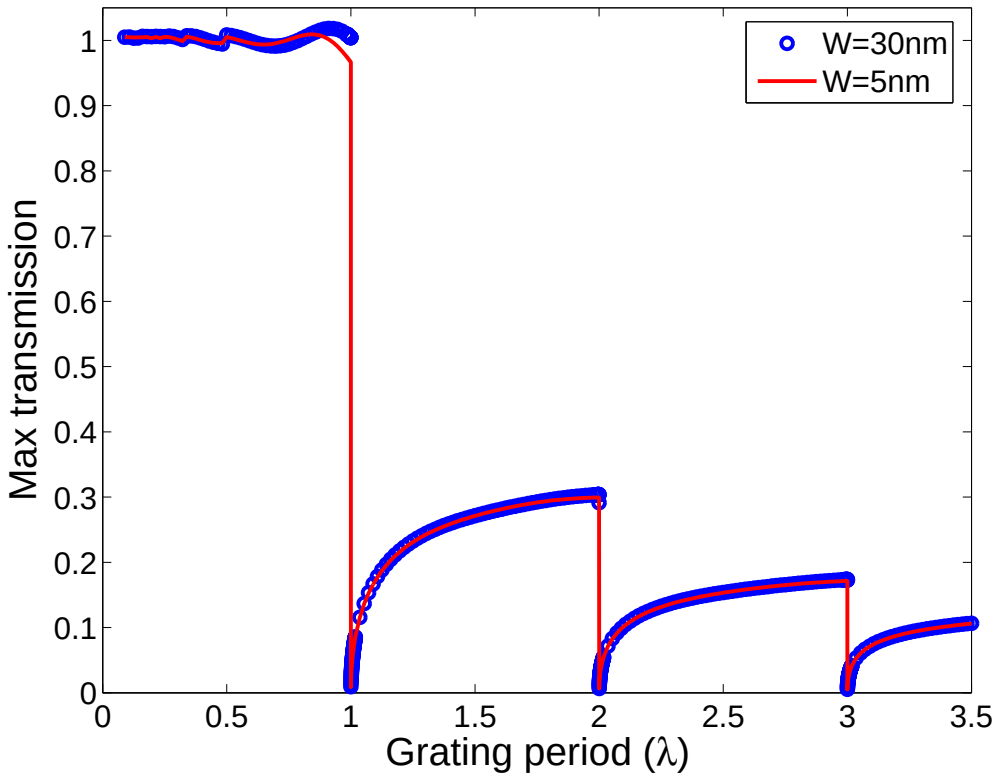

Figure 4.6 The analytic maximum transmission a PEC grating can reach by changing the grating length. The incident wavelength is $\lambda = 560$ nm.

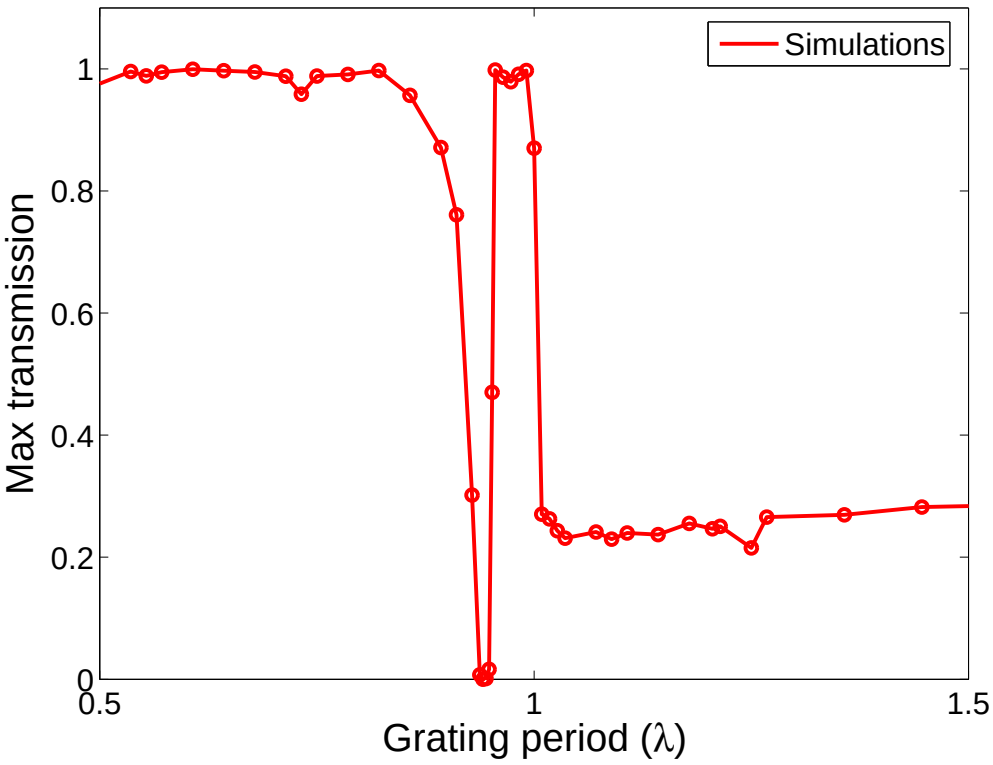

Figure 4.7 The COMSOL Multiphysics simulation results of the maximum transmission a SIBC grating can reach by changing the grating length at different grating periods, The slit width is $W = 30$ nm and the incident wavelength is $\lambda = 560$ nm. The relative permittivity of the metal is $\epsilon_m = -7.92$

transmission of a PEC grating. The analytic curve is calculated by Eq. (4.20). To get the simulation results, I fixed a and W and did many simulations for different d and picked the maximum transmission.

When $a < \lambda$, the maximum transmission is 1 and don't change with a. When $\lambda = a$, the incident wavelength is a resonant wavelength of the surfaces; the transmission is always 0 regardless the grating length, so the maximum transmission is 0, which is verified by the analytic and simulation results in Fig. 4.5.

When $a > \lambda$, the maximum transmission increases with a smoothly and stabilizes at 0.3. $\lambda = a$ is also a Rayleigh wavelength of the grating at normal incidence (70; 71). The Rayleigh wavelengths is the wavelengths that the diffracted wave at another order grazes the surface of the grating. Suppose the medium around the grating is air and the incident plane wave has parallel wavevector $k_{//}$ (parallel with the grating surface), the parallel wavevector of the nth order diffracted wave is $k_{//} + n2\pi/a$. The nth diffracted wave grazes the grating surface when $k_{//} + n2\pi/a = 2\pi/\lambda$, which defines the Rayleigh wavelengths. Obviously $\lambda = a$ is a Rayleigh wavelength at normal incidence. When $a > \lambda$, the reflected fields have more than one propagation modes. More possible channels open for incident energy, so the maximum transmission becomes small.

Fig. 4.6 shows the analytical maximum transmission of PEC grating with different slit widths. When the gratings period increase to a resonant wavelength, the maximum transmission keeps stable. At the resonant wavelength, the maximum transmission jumps sharply to 0. Then it increases smoothly and reach another stable value until the period reaches another resonant wavelengths. The stable values decrease with the increment of the grating period because of the emergence of the new propagation modes.

The most interesting property in Fig. 4.6 is that the maximum transmission doesn't change with the slit width. The curve of $W = 30$ nm overlaps with the curve $W = 5$ nm at any grating period. So we can use very narrow slit to boost the normalized transmission to an arbitrarily big value. So EOT can happen at any wavelengths other than the resonant wavelengths of the input and output surfaces

Fig. 4.7 shows the maximum transmission through a SIBC grating. The resonant wavelengths are different from the Rayleigh wavelength here. At the resonant wavelength, the maximum transmission is 0. At the Rayleigh wavelength, the maximum transmission jumps sharply to a smaller but non-zero

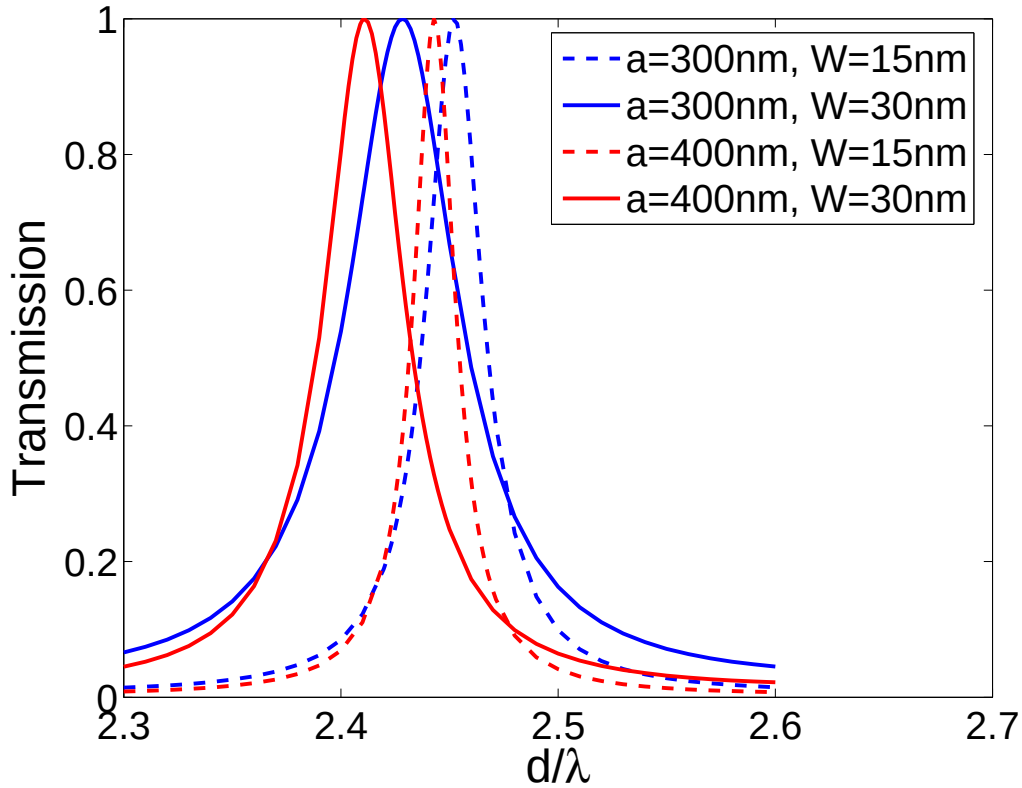

Figure 4.8 Analytic transmission through a PEC grating with a finite length.The incident wavelength is $\lambda = 560$ nm.

value. The difference between the two kinds of wavelengths is very clear here.

4.4.4 Effect of hole shapes

Fig. 4.6 shows the maximum transmission doesn't change with the slit width and when $a < \lambda$, it doesn't change with the grating period too. But the curves of the transmission vs. the grating length do change with a and W. I plot four transmission curves through a PEC grating in Fig. 4.8. Through the height of the transmission peak is 1 for each curve, the locations of the peaks changes with a and W.

Based on Eq. (4.20), the location of the peak is decided by R, and more precisely, the phase of R. R is defined in Eq. (4.13).

$$R = \frac{(\beta_0 + Z_s k_0)\sum_n G_{0n} - P_0}{(\beta_0 - Z_s k_0)\sum_n G_{0n} + P_0} = \frac{k_0 \sum_n G_{0n} - W}{k_0 \sum_n G_{0n} + W}$$

and

$$\sum_n G_n \approx \frac{1}{2\pi}\int_{-W/2}^{W/2}\int_{-W/2}^{W/2} G(y-y')dydy' + \frac{W^2}{\pi}\sum_{n=1}^{\infty} G(na).$$

Here I use a subwavelength PEC grating, so $P_0 = W$, $Z_s = 0$ and $\beta_0 = k_0$.

When I change a and fix W, the only thing changed in the expression of R is $\sum_{n=1}^{\infty} G(na)$. When a is away from the resonant wavelengths, this summation converges and its order of magnitude is around or smaller than 1. When W decreases, both terms in the expression of $\sum_n G_n$ decrease with W. But the second term deceases quickly. So when W is very small, the location of the peak is not sensible to a. Fig. 4.8 verifies it: The dashed lines have smaller W, so the distance between the peaks of the two dashed lines is smaller than the distance of the peaks of the two solid lines.

When W changes, every term in the expressions of R and $\sum_{n=1}^{\infty} G(na)$ changes and they change as different order of W. So the location of the peak changes with W.

4.4.5 Oblique incidence

All the discussion above is based on the normal incidence. Actually the mode expansion method can be extended to oblique incidence easily.

I begin from the general equations about the coefficients of the eigenmodes in the slits of an infinitely-long grating in section 4.2. For the nth eigenmode in the mth slit, we have the equation

$$F_0 I_{mn} - \sum_{m'=-\infty}^{\infty}\sum_{n'=0}^{N-1} A_{m'n'}(\beta_{m'n'} - Z_s k_0)G_{mn,m'n'} = A_{mn}P_{mn}. \tag{4.21}$$

and

$$\begin{aligned}
F_0 &= \frac{2k_\perp}{k_\perp + Z_s k_0};\\
I_{mn} &= \int_{-\infty}^{\infty} \phi^*_{mn}(y)\exp[ik_{//}y]dy;\\
P_{mn} &= \int_{-\infty}^{\infty} \phi^*_{mn}(y)\phi_{mn}(y)dy;\\
G_{mn,m'n'} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dy' \phi^*_{m'n'}(y')\phi_{mn}(y)G(y,y');\\
G(y,y') &= G(y-y') = \int_{-\infty}^{\infty} dk \frac{\exp[ik(y-y')]}{k_x + Z_s k_0}.
\end{aligned}$$

The slits are very narrow comparing with the incident wavelength, so $N = 1$ and $\phi_{m0}(y) = 1$ when y is inside the mth slit. $\Phi_{m0}(x,y) = \phi_{m0}(y)\exp[i\beta_{m0}x]$ is the basic eigenmode in the mth slit. Since all

the slits have the same shape, β_{m0} don't change with m. It is easy to verify P_{m0} is independent with m. So I use β_0 to replace β_{m0}, P_0 to replace P_{m0}. Define $I_{00} = I_0$ and $A_{00} = A_0$.

$$I_{m0} = \int_{ma-W/2}^{ma+W/2} \exp[ik_{/\!/}y]dy = \exp[ik_{/\!/}ma]I_0.$$

and the coefficients of the eigenmodes satisfy $A_{m0} = \exp[ik_{/\!/}ma]A_0$ because of Bloch's theorem. Finally Eq. (4.21) becomes an equation about A_0

$$F_0 I_0 - A_0(\beta_0 - Z_s k_0) \sum_{m'=-\infty}^{\infty} \exp[ik_{/\!/}m'a]G_{00,m'0} = A_0 P_0. \tag{4.22}$$

and

$$A_0 = \frac{F_0 I_0}{P_0 + (\beta_0 - Z_s k_0)\sum_{m=-\infty}^{\infty} \exp[ik_{/\!/}ma]G_{00,m0}}. \tag{4.23}$$

Define $G_{00,m0} = G_m$, then when $m \neq 0$ and $Z_s \neq 0$,

$$G_m \approx \frac{W^2}{2\pi} G(am) \approx \frac{W^2 Z_s i}{h_p} \exp[i2\pi h_p |m| a/\lambda].$$

I neglect the residue quasicylindrical wave term in the Green's function of a metallic surface since its summation always converges.

$$\begin{aligned}
&\sum_{m=-\infty}^{\infty} \exp[ik_{/\!/}ma]G_m \\
&= G_0 + \sum_{m=1}^{\infty} \exp[ik_{/\!/}ma]G_m + \sum_{m=1}^{\infty} \exp[-ik_{/\!/}ma]G_m \\
&\approx G_0 + \frac{W^2 Z_s i}{h_p} \sum_{m=1}^{\infty} \exp[i(k_{/\!/} + 2\pi h_p/\lambda)ma] + \frac{W^2 Z_s i}{h_p} \sum_{m=1}^{\infty} \exp[i(-k_{/\!/} + 2\pi h_p/\lambda)ma]
\end{aligned}$$

The incident angle is θ, so $k_{/\!/} = 2\pi \sin\theta/\lambda$. When $\sin\theta + h_p = n\lambda/a$ (n is an integer), the summation of the second term diverges; when $-\sin\theta + h_p = n\lambda/a$, the third term diverges. So we can define λ_R, the resonant wavelength of a grating surface at oblique incident as

$$\lambda_R = a(h_p \pm \sin\theta)/n; \qquad n \text{ is an integer.}$$

Then the old conclusion is still correct: The transmission is 0 when the incident wavelength is a resonant wavelength.

When the metal is an PEC, $h_p = 1$. We can get the same conclusion by repeating the similar calculations.

We can continue to discuss the transmission through a finitely-long grating and study the locations of the peaks. The maths is exactly the same and all the basic conclusions still hold. So I neglect further discussion here.

4.4.6 Conclusion

In this section we developed a theory about extraordinary transmission through a subwavelength metallic grating from the first principles using mode expansion method. The theory includes three parts: the Green's function expression of surface waves; multiple scattering of surface waves and the Fabry-Pérot interference between input and output surfaces. The basic conclusions include

1. The transmission has dips at the resonant wavelengths of the input and output surfaces.
2. EOT can happen at any wavelengths other than the resonant wavelengths.
3. Surface waves connects the slits in the grating and play important roles at both transmission dips and peaks.
4. The transmission peaks result from the Fabry-Pérot interference between the input and output surfaces.
5. The peak wavelengths are decided by the details of the structures, such as the the grating period, the slit width and the grating length.

4.5 Application of the extraordinary transmission theory

In the previous section I present a EOT theory. Now I will use my theory to explain the figures from others' papers.

In the previous section I always fix the incident wavelength plot the transmission versus a geometry parameter. It is convenient for analytic calculations and it makes the physics underneath clear. In my theory the slits are connected by the surface waves, which is decided by the surface parameters such as a and W; the input and output surface are connected by the eigenmodes in the slit, which are decided by W and d. So It is necessary to study the geometry parameters separately.

However experimentalists prefer keeping the grating unchanged and running the wavelength scan. Scaling law says that changing the wavelength is equivalent with fixing the wavelength and changing all the geometry parameters proportionally, if we neglect the wavelength dependence of the relative permittivity of the metal. So the surface resonance and the Fabry-Pérot resonance mix inevitably.

4.5.1 Phys. Rev. Lett. **83**, 2845 (1999) by J. A. Porto *el al.*

The paper "Transmission Resonances on Metallic Gratings with Very Narrow Slits" by J. A. Porto, F. J. García-Vidal and J. B. Pendry presented a theoretical analysis about EOT through a metallic grating (66). They calculated the transmission spectrum by using transfer matrix formalism based on the mode expansion method. The key conclusion of he paper is that there are two kinds of transmission peaks: the peak coming from the the excitation of coupled surface plasmon on both surfaces of the grating and the peak coming from the excitation of waveguide resonances located in the slits.

Fig. 4.9 is excerpted from the the paper. The top part shows the grating under study. The authors used different notations. In their paper, d is the period of the grating; a is the width of the slits; h is the length of the grating or the thickness of the metallic film. The bottom part shows the transmission spectra when the grating period is 3.5 μm, slit width is 0.5 μm. Their study focused on the transmission peaks. For example, when the thickness of the film is 3 μm, the peak at $\lambda \approx 4$ μm comes from the coupled surface plasmon resonance and the the peak at $\lambda \approx 7.5$ μm comes from the waveguide resonance.

Now I will use my theory to explain the transmission spectra. First though the locations of the peaks change with the thickness of the film, the locations of the dips don't change. We can find dips at $\lambda \approx 3.5$ μm and $\lambda \approx 1.8$ μm, which are the resonant wavelengths of the input and output surfaces.

Second let's focus on the peaks. As I have discussed, the peaks appear when $R\exp[i\beta_0 h]$ is a real number. When λ changes, both R and β_0 changes. β_0 always changes smoothly with the wavelength. We can assume $\exp[i\beta_0 h]$ is a constant in a narrow region around the resonant wavelengths. But because of the resonance, R changes rapidly in this region and cover a broad phase interval. So for most of h value, the peak condition will be fulfilled around the resonant wavelength. So we can always find a peak around the resonant wavelength. But it is still possible that no peak appears around the resonant

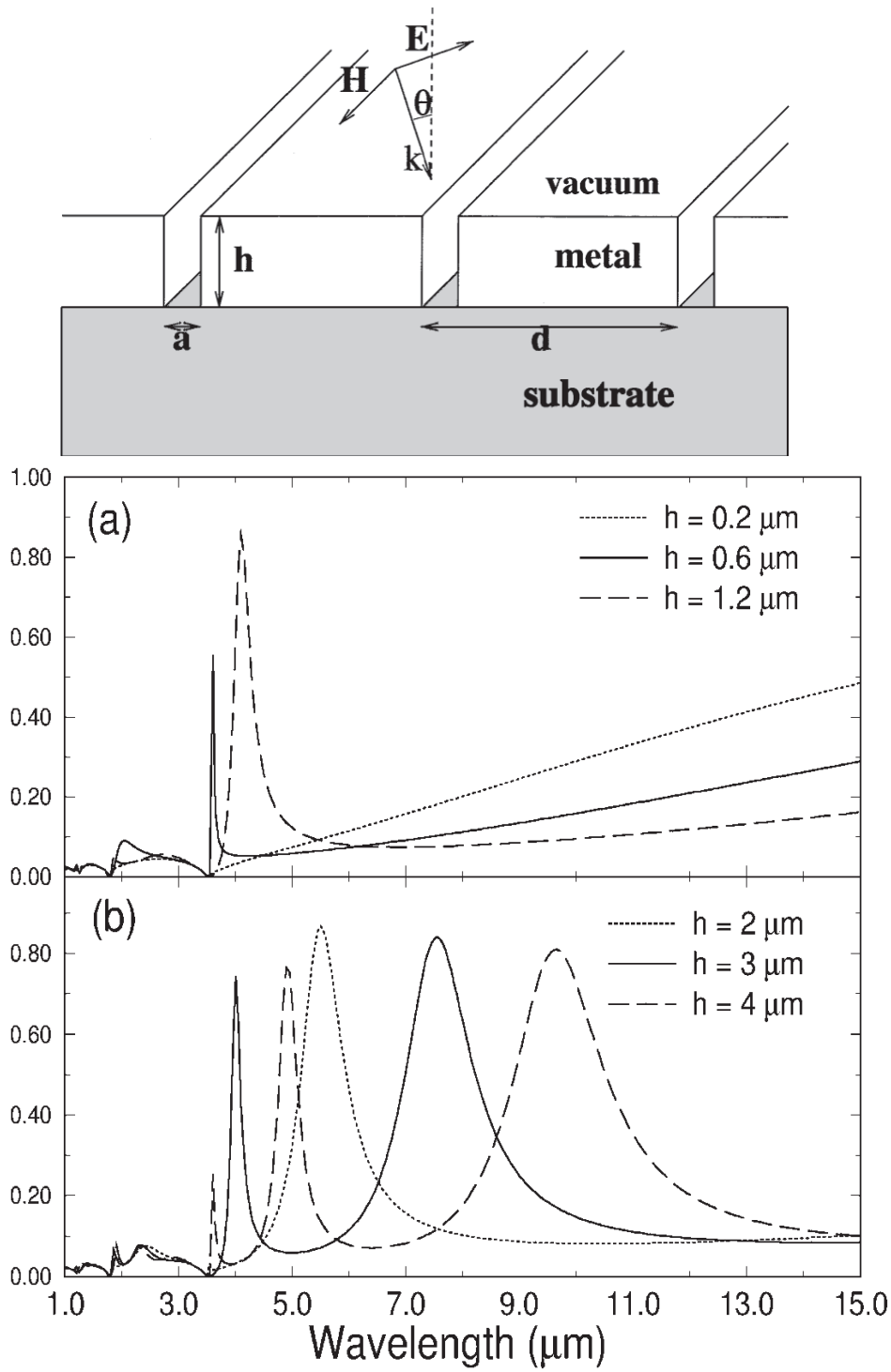

Figure 4.9 Figures excerpted from PRL **83**, 2845 (1999). (Top) Schematic representation of the metallic grating studied in the paper. (Below) Zero-order transmission through the grating vacuum at normal incidence for different values of the grating length. The grating period is $d = 3.5\ \mu$m and the slit width is $a = 0.5\ \mu$m.

wavelength since the phase of R don't cover the whole 2π interval. When $h = 0.2\ \mu$m and $2\ \mu$m, there is no peak around the resonant wavelength. When h is big, the phase interval covered by $\exp[i\beta_0 h]$ becomes bigger, so more peaks appear at the wavelengths away from the resonant wavelength.

Around the resonant wavelength, the phase of R changes very quickly with the wavelength, so the linewidths of the peaks are very narrow. Away form resonant wavelength, both R and β_0 changes smoothly, so the linewidths of the peaks are wide.

The heights of the peaks also agree with my theory. The peak height increases when the peak wavelength increases from the the resonant wavelength. Actually the peak reaches its full height quickly. The transmission don't reach 100% here is because the metal is lossy in this paper. When λ is smaller than the grating period, the peaks become low.

In conclusion, the simulation results presented in this paper agree with my theory perfectly.

4.5.2 Phys. Rev. Lett. **92**, 183901 (2004) by K. J. Klein Koerkamp *el al.*

The paper "Strong Influence of Hole Shape on Extraordinary Transmission through Periodic Arrays of Subwavelength Holes" by K. J. Klein Koerkamp, S. Enoch, F. B. Segerink, N. F. van Hulst and L. Kuipers discussed the effect of hole shapes (72). The authors measured the transmission through the periodic subwavelength hole arrays in a gold film. They verified, experimentally and numerically, that the shape of the holes influence not only the heights but also the central wavelengths of the normalized transmission peaks.

The top figures in Fig. 4.10 shows the focused ion beam images of two periodic subwavelength hole arrays used in the experiments. The bottom figure of Fig. 4.10 is the normalized transmission spectra calculated by the Fourier modal method. Three hole shapes were studied: circular holes with a diameter of 190 nm, rectangles holes of 150 nm $\times$ 225 nm and of 75 nm $\times$ 225 nm. The lattice constant is 452 nm $\times$ 425 nm. The films were deposited on glass substrate.

I have analyzed the effect of hole shapes analytically, so the main conclusion about the strong influence of holes shapes presented in this paper is not very surprising. Now let's analyze the normalized transmission curves in Fig. 4.10 quantitatively. Again, we find the transmission dips don't change with

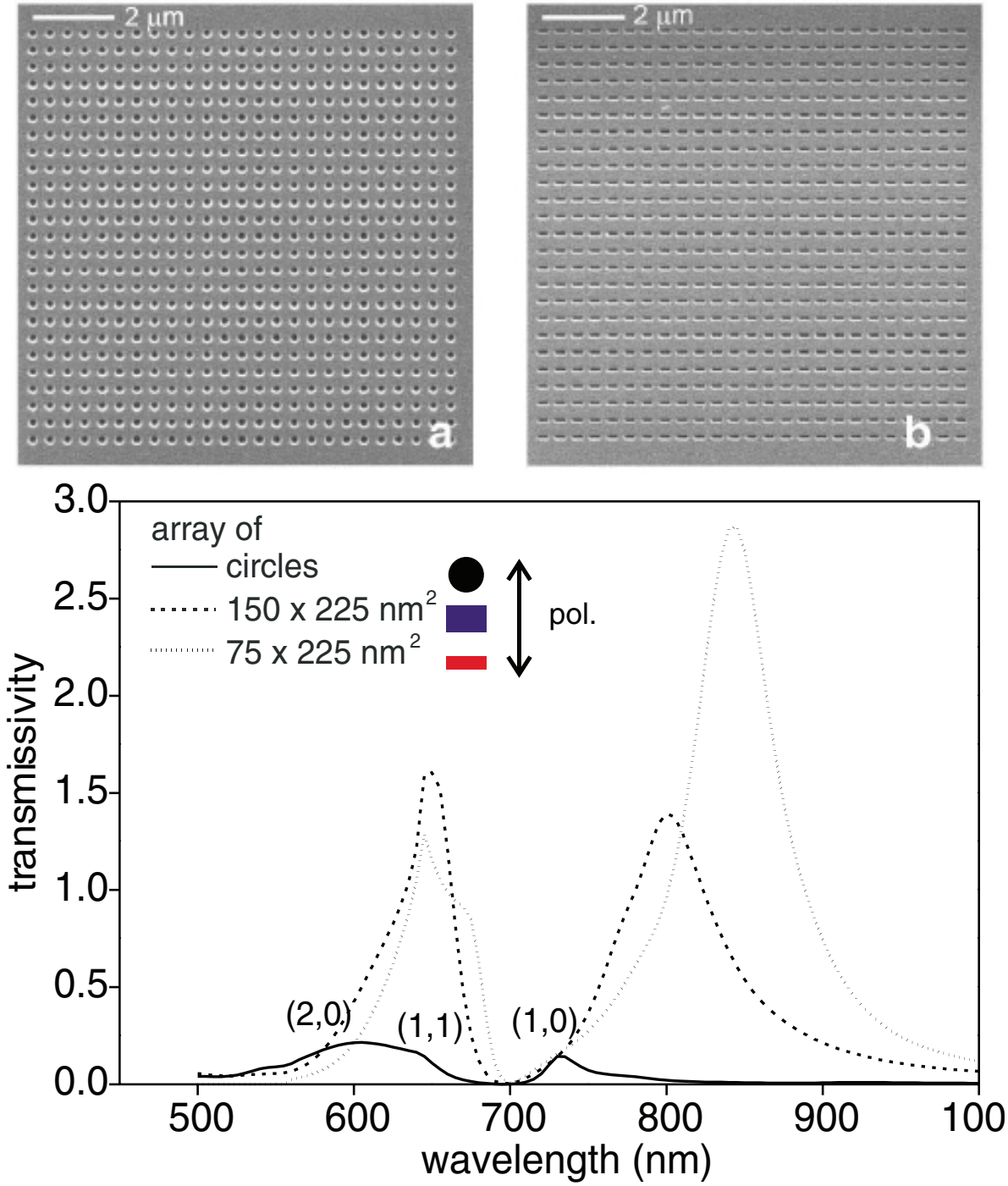

Figure 4.10 Figures excerpted from PRL **92**, 183901 (2004). (Top) Images of two periodic hole arrays used. Both arrays have period d = 425 nm. (a) Circular holes with a diameter of 190 nm (b) Rectangular holes of 75 nm × 225 nm. (Below) Calculated normalized transmission of three different subwavelength periodic hole arrays. E field of the incident wave was perpendicular to the long axis of the rectangles.

the hole shapes. Since 2D hole array is used here instead of the 1D slits, the resonant wavelengths are

$$\lambda_R = \frac{na}{\sqrt{i^2 + j^2}} \sqrt{1 - Z_s^2}.$$

Here λ_R is the wavelength in free space; n is the refractive index of the substrate; a is the lattice constant since the hole array has a square lattice. The first dip is at $\lambda \approx 700$ nm. So the second dip should be around 700 nm/ $\sqrt{2}$ = 495 nm and we can find this dip in Fig. 4.10 too.

The height of the normalized transmission peaks doubles when the area of the rectangle holes is halved. So the heights of the transmission peaks equal each other for the rectangle hole shapes, which agrees with my theory. The height of the peaks of the circular holes is very low. The authors of the paper contributed it to the absence of the shape resonances in the circular holes. Based on my theory, another possible explanation is that the peak is too close to the resonance wavelengths of the surfaces. Because no propagation modes exist in the holes, I can not assert which explanation is more reasonable. Further study is necessary here. My theory cannot explain the heights of the peaks between 500 nm and 700 nm.

In conclusion, the experimental and numerical work on the EOT through 2D subwavelength holes in this paper confirms (1) the transmission is low at the resonant wavelengths of the input and output surfaces and (2) the hole shapes influence the transmission peaks substantially in a complex way.

4.5.3 Phys. Rev. Lett. **96**, 233901 (2006) by Z. Ruan and M. Qiu

The paper "Enhanced Transmission through Periodic Arrays of Subwavelength Holes: The Role of Localized Waveguide Resonances" by Z. Ruan and M. Qiu analyzed the transmission through a square lattice of rectangle air holes in a free standing PEC film, which is shown in Fig. 4.11 (73). In this paper, the holes are always $0.9a \times 0.2a$ and the lattice constant d has three possible values: $1.0a$,$1.1a$ and $1.2a$. The thickness of the film is $h = 0.2a$. Here a is an arbitrary length unit. The incoming wave is normal incident plane wave with electric field along the short edge of the rectangular holes. The normalized transmission was calculated by the full-vectorial 3D finite-difference time-domain method and is shown in Fig. 4.12.

As Fig. 4.11 shows, the rectangle holes are narrow along y direction and long along x direction, so the open area is quite close to 2D slit. The cutoff wavelength of the basic mode in the hole is $1.8a$ and

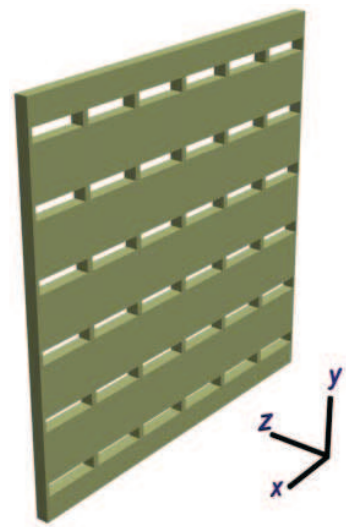

Figure 4.11 A figure excerpted from PRL **96**, 233901 (2006). Schematic representation of a PEC film with a periodic square array consisting of rectangular air holes studied in the paper.

the cutoff frequency is $0.556(c/a)$. So there exists propagate mode in the mode at resonant frequencies of the surfaces, which is also similar to the case of slits. Then it is not surprising that they got the similar normalized transmission curves at resonant frequencies. We can find a dip and a peak there and their location are decided by the lattice constants. We can also find peaks away from the resonant wavelengths and their locations are not sensitive to the lattice constants. Through the heights of the normalized transmission peaks are different, the heights of the transmission peaks are all close to 1. All these facts agree with my theory.

We should also the peaks away from the resonant frequencies are close to the cutoff frequencies. This can also be explain be my theory. The peak condition is $R\exp[i\beta_0 h]$ is real. When the wavelength is far away the resonant wavelengths and $A \ll \lambda^2$ (here A is the area of the hole), $R \approx -1$. At the cutoff wavelength, $\beta_0 = 0$,. so $R\exp[i\beta_0 h] \approx -1$ is a real number. Then we get a peak. When the long edge of the rectangle hole becomes longer, the hole converges to a slit, the cutoff wavelength becomes longer too, so the peak appear at longer wavelength and the transmission will increase monotonically when the wavelength is not so large. The curves in Fig. 4.9 show the monotonic increment when $h = 0.2, 0.6, 1.2\ \mu$m.

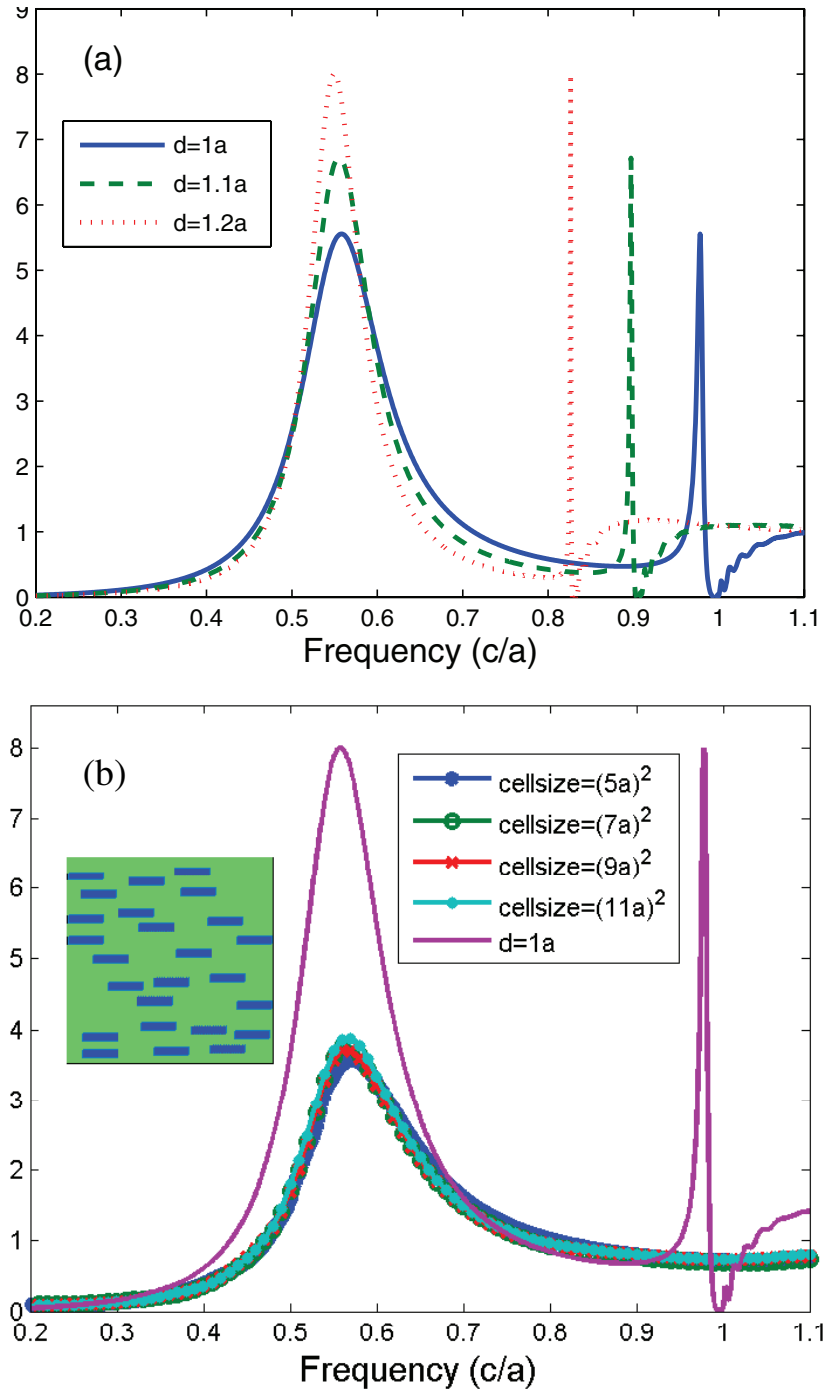

Figure 4.12 Figures from excerpted PRL **96**, 233901 (2006). (Top) Normalized transmission through a PEC film with periodic hole arrays for different lattice constants d. The thickness of the film is $0.2a$. The size of the hole is $0.9a \times 0.2a$. (Below) The mean value of the normalized transmission of the random sample with different cell size. The solid purple line shows the normalized transmission of the periodic hole array with the lattice constant $d = 1.0a$. The hole size of the random sample is still $0.9a \times 0.2a$. The filling ratio of the holes equal to that of the periodic array. The inset shows a example of a sample with 25 randomly distributed holes in a $5a \times 5a$ cell.

This paper also studied the transmission when the holes are randomly distributed. The normalized transmission curves are shown in the bottom of Fig. 4.12. Comparing with the curve of the periodic holes, the peak and the dip at resonant frequency disappear and the peak at the cutoff frequency remains. This result is very easy to understand. When the holes distribute randomly, the summation of Green's function terms will not go to infinity. So the surface wave resonance disappears. When the area of the holes is small, R is always close to 1, so the peak appear at the cutoff frequency only.

I have to emphasize here though the transmission peaks come from the Fabry-Pérot resonance of the input and output surfaces, it doesn't mean the surface waves are not important when the peaks are away from the resonant wavelengths. Physically, surface waves collect the energy hitting the surface and bring it to the holes. Mathematically, the maximum transmission is decided by T and R which has strong relations with the surface waves. So surface wave is important. On the other side, the peak close to the resonant wavelength also comes from the Fabry-Pérot resonance of the input and output surfaces. The surface wave resonance will always bring a transmission dip.

4.5.4 Conclusion

In this section I used my EOT theory to explain the simulation results shown in literatures successfully. In the future I like to extend my theory to 2D periodic holes cases. As the second paper I analyzed in the section shows, the 2D periodic hole array will bring us new physics.

CHAPTER 5. Beaming

5.1 Introduction of beaming

5.2 Beaming of two-layer dielectric rods

In the first chapter I have illuminated the principle of beaming using a photonic crystal with a surface layer and a grating layer. However we have discovered that the photonic crystal is not necessary for surface modes and a single dielectric layer support surface mode too (39). So theoretically a similar beaming phenomenon will happen for a two-layer structure: one surface layer to support the surface modes and one grating layer to couple the surface modes to radiation modes. In this section, I demonstrate the converging property of the two-layer structure. Furthermore, by putting several two-layer structure in series, we can sustain the beaming to long distances.

5.2.1 A two-layer structure

The 2D two-layer structure is shown in Fig. 5.1(a). The top layer is the surface layer. It consists of 41 circular rods with diameter $D = 1.83$ mm and lattice constant $a = 11$ mm. The bottom layer is the grating layer with 21 square rods with the side length $L = 3.15$ mm; the lattice constant is $b = 2a = 22$ mm. The distance between the two layers is a. All the rods are made of alumina and the permittivity is 9.8.

The commercial finite element method software COMSOL Multiphysics was used to simulate the system. The simulation area with n structures is shown in Fig. 5.1(b). A Gaussian beam with TM polarization (the electric field parallel to the dielectric rods) is incident normally from the left boundary. The waist of the Gaussian beam is 3λ. The power flow through the red lines given in Fig. 5.1(b) is calculated by integrating the time average normal component of the Poynting vector along the lines.

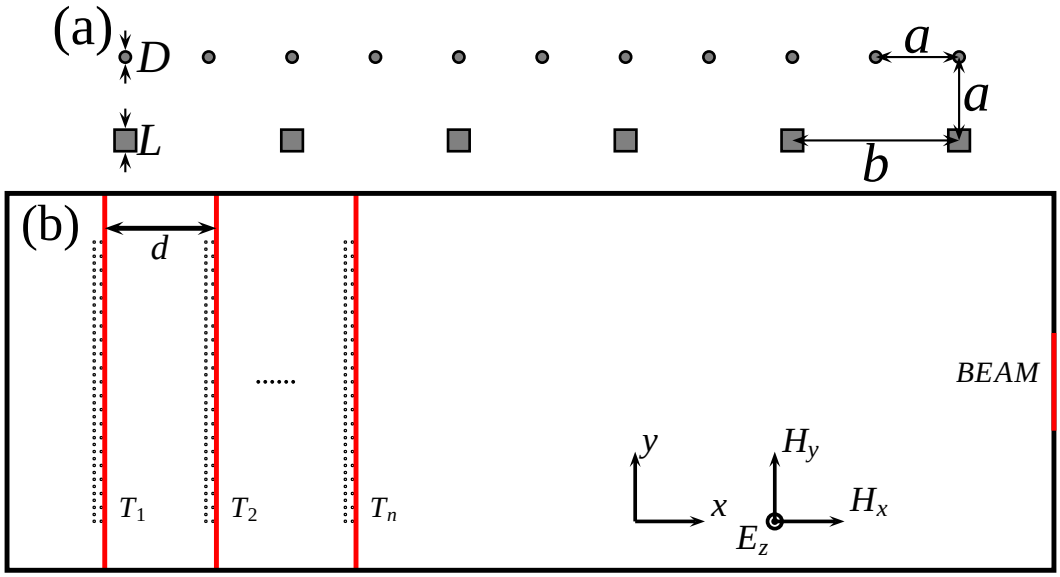

Figure 5.1 (a) A schematic drawing of the two-layer structure and (b) a schematic drawing of the simulation area.

The power flow through the red line right after the ith structure is defined as T_i and the power flow through the short red line on the right boundary as $BEAM$ (see Fig. 5.1(b)). The length of the short red line is $14a$. Suppose the total input power is I. T_n/I describes the transmission after the n structures. Then we define the distribution factor as

$$\text{Distribution factor} = \frac{BEAM}{T_n}.$$

The distribution factor definition is just a simple quantitative index of the directionality of the output field through the n structures. The space between the last long red line and the right boundary, shown in Fig. 5.1(b), is $99a \times 54a$ for all the simulations. This way we can compare the distribution factors between different simulations.

First the converging ability of a single two-layer structure is studied. We compare four cases: empty simulation area, a surface layer only, a grating layer only and a two-layer structure. The simulation results are shown in Fig. 5.2 when incident frequency is between 9 GHz to 13 GHz. The transmission and the distribution factor curves are shown on the left column. When the frequency is above 13 GHz, the transmission field has more than one beams, which is against my wish to improve the directionality. When the frequency is below 9 GHz, the distribution factor curves after the rods converge to the empty

area curve.In Fig. 5.2(a), the simulation area is empty, so the transmission is one. In Fig. 5.2(c), the surface layer decreases the transmission a little bit but the distribution factor curve overlaps the empty area curve. In Fig. 5.2(e), the distribution factor curve of the grating layer is close to the empty area curve when the frequency is small. It has a dip when frequency is around 12.7 GHz. The corresponding wavelength is 23.6 mm, which is close to $b = 22$ mm, the period of the grating layer. So the interaction between the field and the rods is strong. But this interaction don't improve the directionality of the transmission field. The two-layer structure has a much different distribution factor curve (Fig. 5.2(g)). The curve has two dips which are transmission dips too. The interesting frequency region is around 11.5 GHz. The transmission fields in this frequency region have big distribution factors. When $f = 11.59$ GHz, the distribution factor reach a maximum value of 0.863. The transmission is acceptable too, it's around 0.87.

To understand how the directionality is improved, the electric field along the z-axis, E_z, at the red line after the rods (see Fig. 5.1(b)) at $f = 11.59$ GHz is plotted on the right column of Fig. 5.2. The rods of the surface layer are located at $y = -20a, -19a, \cdots, 20a$ and of the grating layer are $y = -20a, -18a, \cdots, 20a$. The field on the right column of Fig. 5.2 is basically the superposition of the transmission Gaussian beam and some quick oscillations. The field after the surface layer has small oscillation. It means the interaction is weak. The oscillation is the surface state which decays exponentially along x direction. Though the existence of the surface layer reduces the transmission a little bit, it cannot change the distribution factor (see Fig. 5.2(c)). Fig. 5.2(f) and Fig. 5.2(h) shows the strong interaction between the field and the rods. Comparing the Gaussian envelopes of the fields in the two figures, we can find the envelope in Fig. 5.2(f) has the same width as the incident beam and Fig. 5.2(h) is wider. It is because the surface layer of the two-layer structure supports the surface states. The surface states will propagate along the surface layer and results in wider field distribution. This wider field distribution will bring us better directionality.

The electric field distributions of the empty simulation area and the simulation area with the two-layer structure at $f = 11.59$ GHz are shown in fig. 5.3. Fig.5.3(a) shows a typical 2D Gaussian beam in free space. The beam waist is at the the input surface and the width of the beam grows during the propagation. Comparing with Fig. 5.3(a), we can find that the width of the beam in fig.5.3(b)

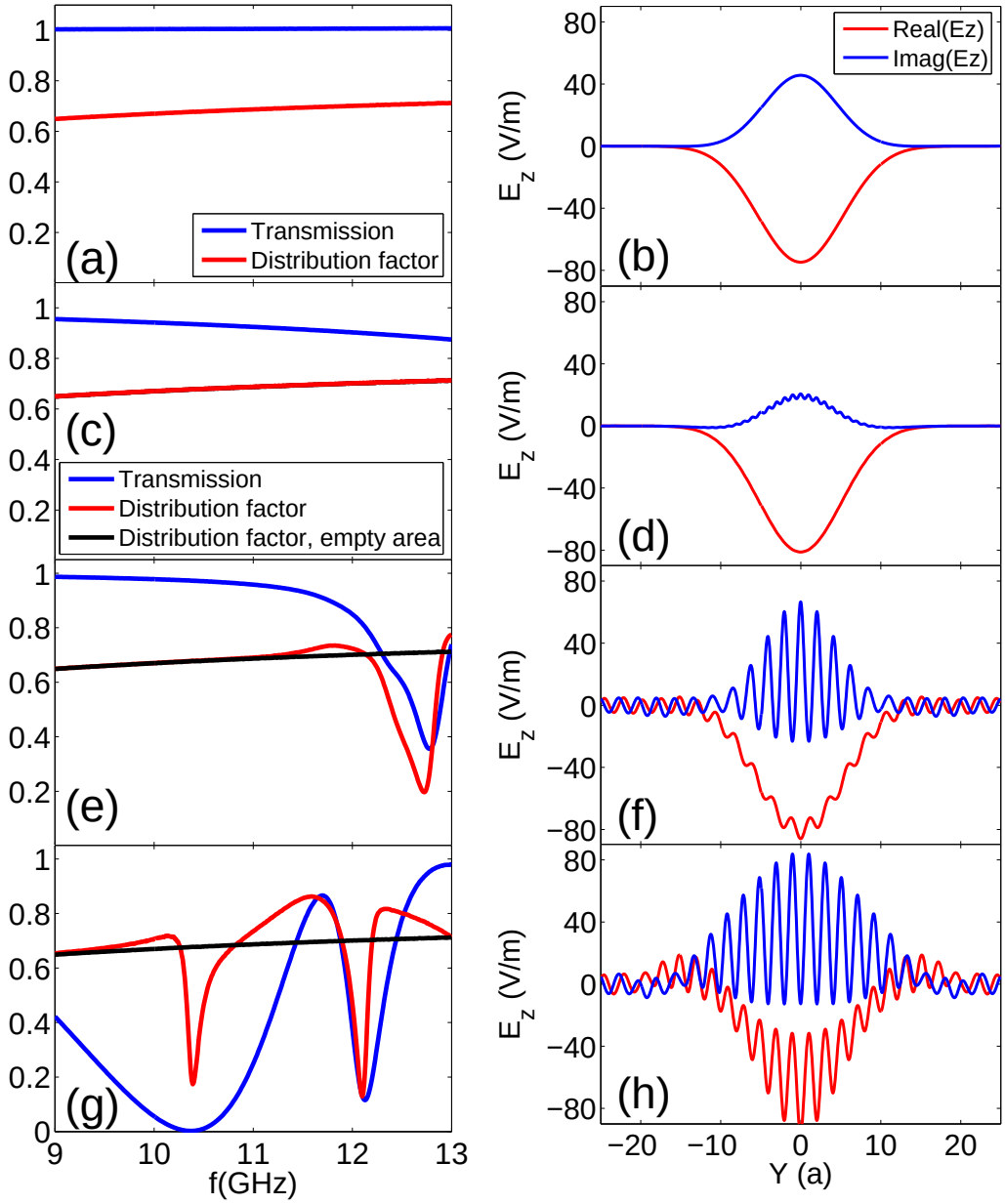

Figure 5.2 Simulation results of empty area (a, b), a surface layer only (c, d), a grating layer only (e, f) and a two-layer structure (g, h). The left column shows how the transmission and the distribution factor change with frequency. The right column shows the electric field along the red line right after the rods when f = 11.59 GHz.

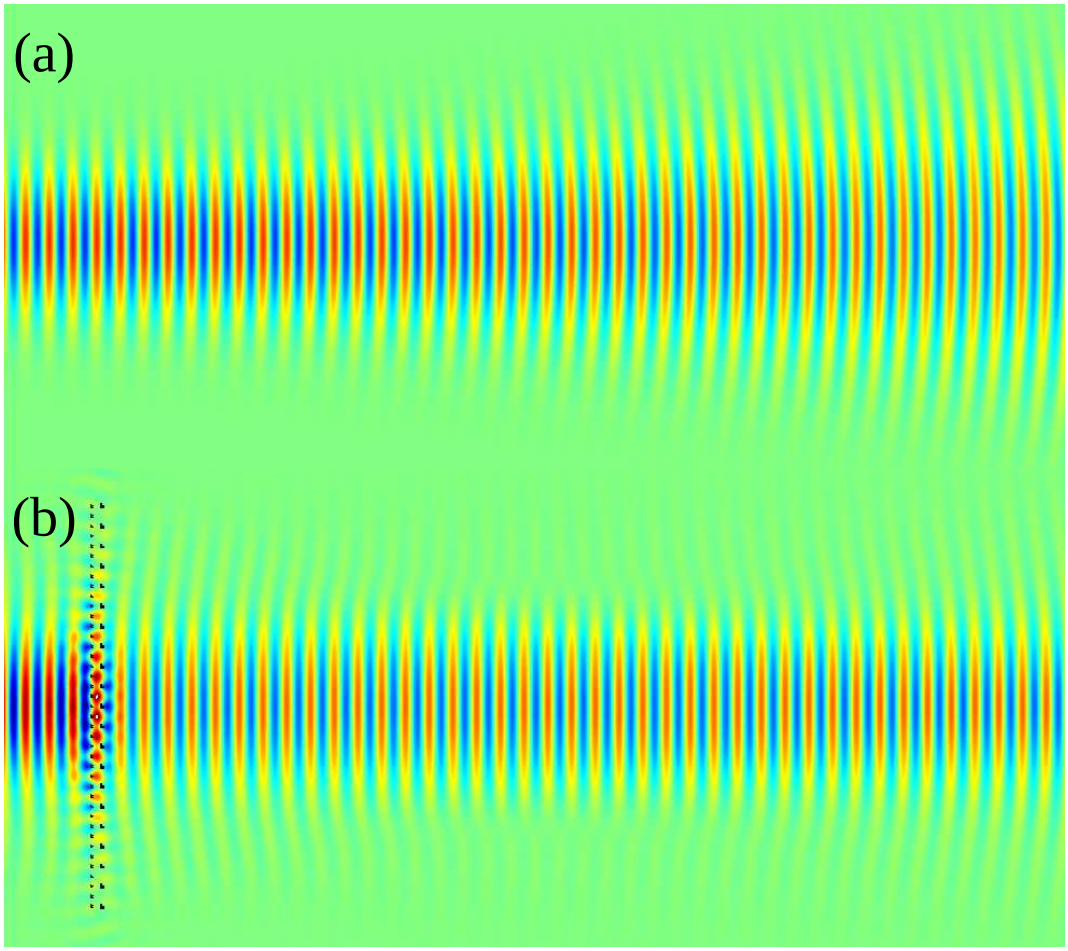

Figure 5.3 E_z distribution of the simulation area (a) without and (b) with a two-layer structure when f = 11.59 GHz.

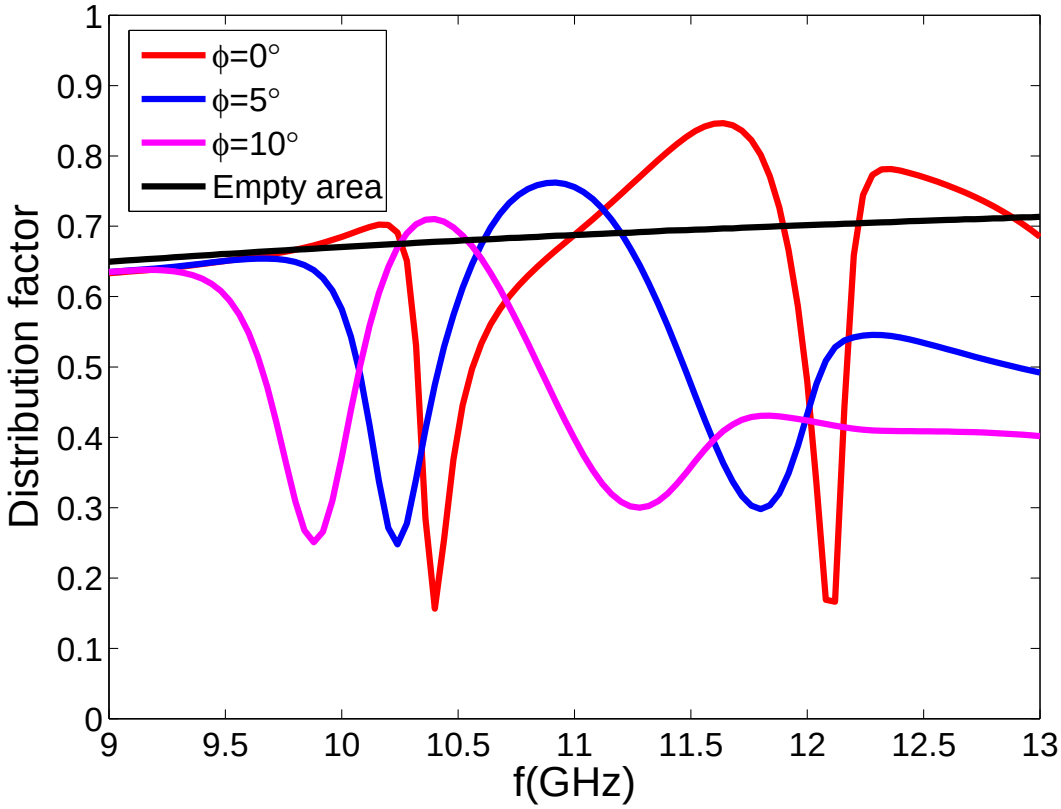

Figure 5.4 Distribution factors of fields after the two-layer structure illuminated by by oblique Gaussian beams. ϕ is the incident angle.

almost doesn't change during the propagation. We can also see the surface mode along the structure in Fig. 5.3(b). The figure shows the converging ability of the two-layer structure clearly. The converging ability does not seems very strong here. It is because that the Gaussian beam has good directionality already.

The converging of oblique incident Gaussian beams is also studied. The distribution factors of the fields after the two-layer structure when the incident angle ϕ equals $0°, 5°, 10°$ are shown in Fig. 5.4. The two-layer structure improves the directionality of transmitted field when the incident angle is small. The greatest improvement happens at normal incidence. The maximum distribution factors appear at f =11.64 GHz, 10.92 GHz and 10.40 GHz respectively. The band structure of the surface layer is shown in Fig. 5.5. When we fix the incident angle and scan the frequencies, the wave vector component of the incident wave parallel with the two-layer structure is $2\pi f \sin\phi/c$; here c is the speed of light in free space. Considering the function of the grating layer, the wavevectors of the surface modes which can

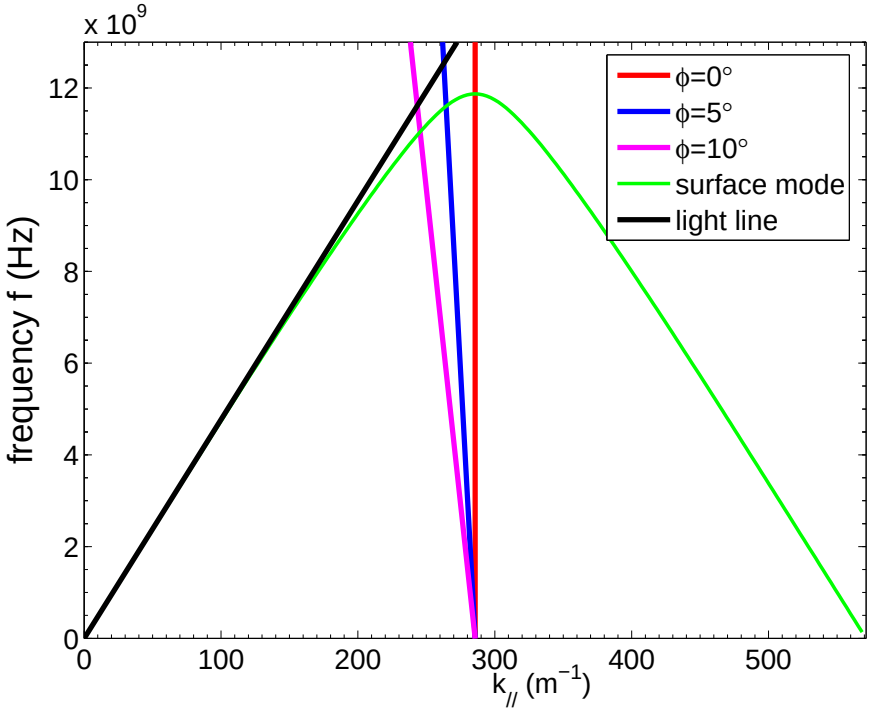

Figure 5.5 Band structure of the surface layer along with the wavevectors of the surface modes which can be excited by the incident waves when the incident angles are $0°, 5°$ and $10°$.

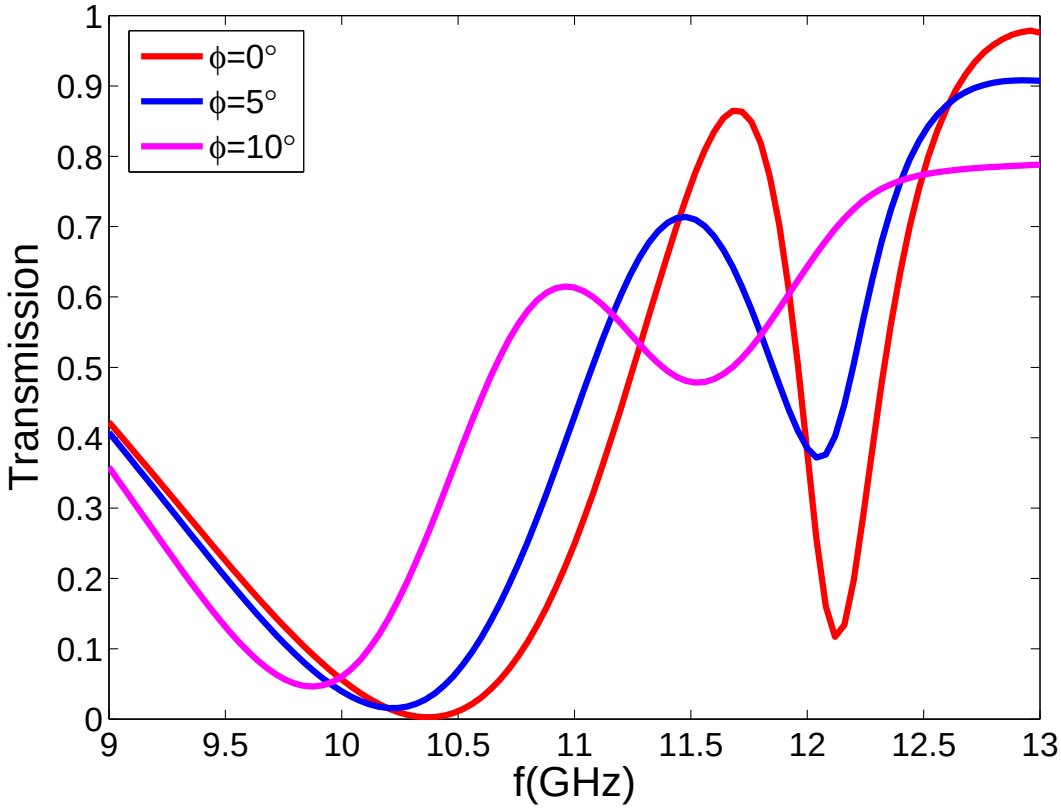

Figure 5.6 Transmission of oblique Gaussian beams through the two-layer structure.

be excited by the incident wave are $k_{//} = \pi/a - 2\pi f \sin\phi/c$. This relation at the three incident angles is also plotted in Fig. 5.5. The frequencies of the crossing points are 11.97 GHz, 11.60 GHz and 11.06 GHz when $\phi = 0°, 5°, 10°$. These frequencies decease when the incident angles increase, same as the maximum distribution factor frequencies. It supports the explanation that surface modes improve the directionality. Actually the gaps between the two frequencies are around 0.3 GHz at the three incident angles. We can contribute the gap to the influence of the grating layer; the band structure of the surface layer is calculated without the grating layer.

From Fig. 5.4 we can find the maximum distribution factors decrease with the increasing incident angles. The band structure of the surface layer provided a possible explanation. As Fig. 5.5 shows, when incident angle increases, the crossing points come closer to the light line. It means that the surface mode become flat along its decaying direction and difficult to be excited strongly. So the distribution factors decease.

Fig. 5.4 also shows the distribution factors become small then the frequencies are big. It is because more than one propagation modes existing after the two-layer structure when frequency is big. Given incident angle ϕ, the second propagation mode emerges when

$$f = \frac{c}{2a}\frac{1}{1+\sin\phi}.$$

The critic frequencies at the three incident angle 0°, 5°, 10° are 13.6 GHz, 12.5 GHz and 11.6 GHz respectively, which agree with the Fig. 5.4 quite well.

The transmission of the oblique Gaussian beam through the two-layer structure is shown in Fig. 5.6. The transmission peak overlaps with the distribution factor peak only when $\phi = 0°$.

5.2.2 More two-layer structures

We have understood the converging ability of a two-layer structure. Now I begin to examine the two two-layer structures in the simulation area. The working frequency is $f = 11.59$ GHz, the optimal converging frequency of the single structure. Fig. 5.7 shows the transmission and the distribution factor of the field as a function of the distance between the two structures d. The transmission curve is periodic with a $\lambda/2$ period. The curve can be explained easily by the one mode assumption. Let's consider the simplest case: the incidence of a Gaussian beam to a single structure. Suppose the incidence, reflection

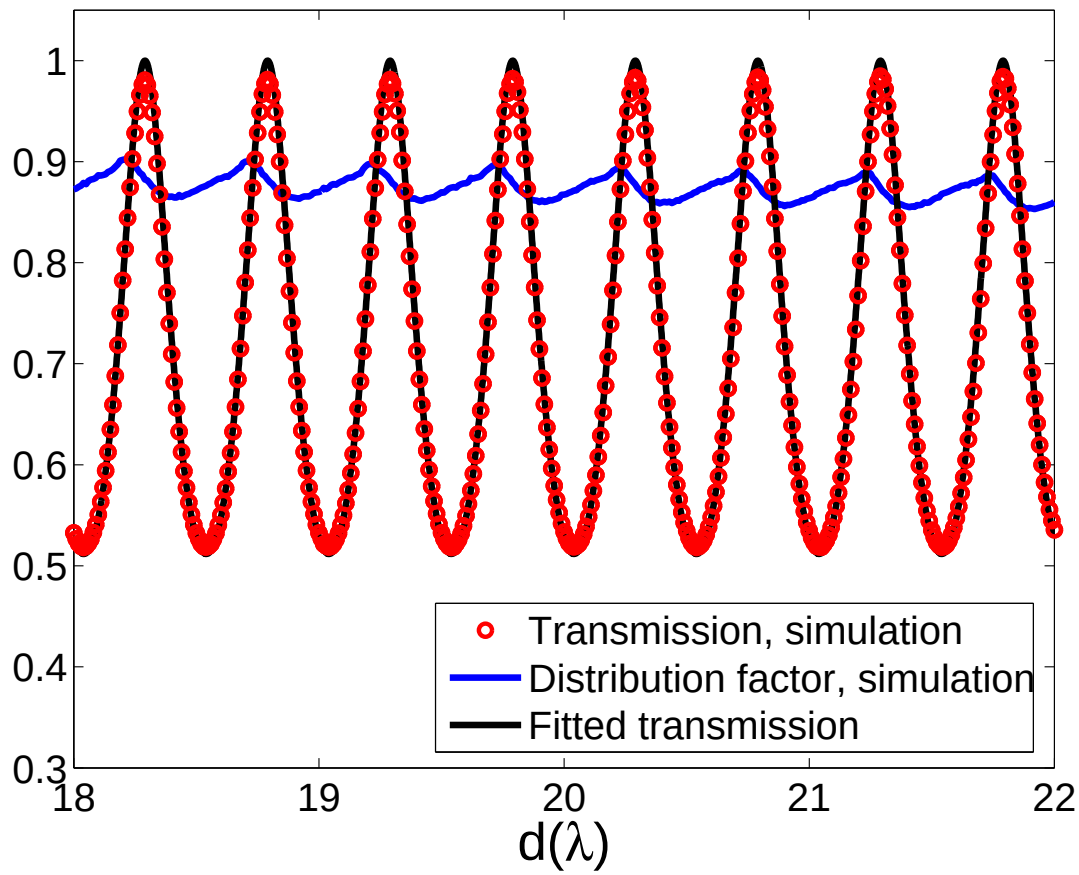

Figure 5.7 Transmission and distribution factor of the field after two two-layer structures. The black line shows the fitted transmission using the equation of the transmission matrix theory.

and transmission electric fields are $\Phi_i(x, y)$, $\Phi_r(x, y)$ and $\Phi_t(x, y)$. The one mode assumption says that $\Phi_i(x, y) \approx \phi(y)\exp(ikx)$ and $\Phi_r(x, y) \approx A\phi(y)\exp(-ikx)$, $\Phi_t(x, y) \approx B\phi(y)\exp(ikx)$. Here $k = 2\pi f/c$ (c is the speed of light), the reflection A and transmission B coefficients are complex numbers. Only one mode exists during the propagation, reflection and transmission. In fact, the Gaussian beam will diverge slowly during the propagation. But if the free length of space is short, like 20λ, the divergence is small and the lateral profile don't change too much. So $\Phi_i(x, y) \approx \phi(y)\exp(ikx)$. Then the Gaussian beam hits the structure. The structure converges the beam and cancels the diverging trend during the propagation. So the reflection and transmission fields have a similar profile.

Under the one mode assumption the transmission matrix of the system is a 2×2 matrix. Suppose

the transmission matrix of one two-layer structure is M and the propagation transmission matrix is P,

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}; P = \begin{pmatrix} \exp(ikd) & 0 \\ 0 & \exp(-ikd) \end{pmatrix}.$$

Then the transmission T after the two structures is (74)

$$T = \left| \frac{1}{|m_{11}|^2 - 1 + m_{11}^2 \exp(2ikd)} \right|^2 \tag{5.1}$$

Use this equation to fit the simulation transmission, we get $m_{11} = 1.058 - 0.279i$. The fitted transmission is also shown in Fig. 5.7.

The distribution factor curve in Fig. 5.7 is a periodic function too. The interesting property of the curve is that its oscillation amplitude is small. This is a natural result of the one mode assumption. The field after the last structure is always proportional to $\phi(y)\exp(ikx)$. So the distribution factor shouldn't change as we increase the distance d.

When $d = 18.29\lambda, 18.79\lambda, 19.29\lambda, \cdots$, the transmission is close to 1. So the transmission field is $\Phi_t(x, y) \approx B\phi(y)\exp(ikx)$ and $|B| \approx 1$. The only difference between the transmission field and the incident field is a phase factor. The two structures are transparent to the incident field. Then we can put more structures after the transmission beam to repeat the pattern. A single mode waveguide for the Gaussian beam is formed by the equidistant two-layer structures. Fig. 5.8 shows the time average energy density distributions of the waveguide. We can see clearly that the width of the beam does not change during the propagation. After 8 structures (which corresponds to distance of 191λ) , the transmission is 0.930 and the distribution factor is 0.874. Some power is lost because of the surface mode but most of the power is guided.

5.2.3 Conclusion

In previous studies, beaming is always referred to the directionality of the output field of a channel. The channel could be a hole in a metal film(46), or a subwavelength metal slit(53; 54; 55; 56), or a line defect in a 2D photonic crystal(48; 49; 50; 51; 52). In fact, beaming can have broader meaning based on the source. Our design is a good example of Gaussian beam beaming; others' designs work for channel modes. The beaming light from a subwavelength channel is similar to a Gaussian beam(55; 48). The

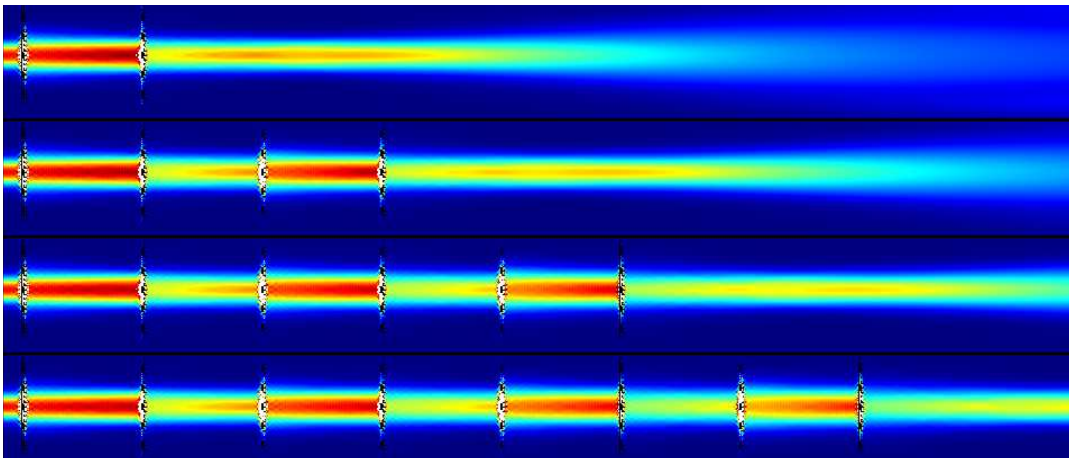

Figure 5.8 Time average energy density distribution when the simulation area has 2,4,6 and 8 two-layer structure. The frequency is $f = 11.59$ GHz and the distance between two adjacent structure is $d = 21.29\lambda$.

two-layer structure provides a method to extend the beaming for long distances. Finding an effective beaming device for a point source is an interesting problem.

Another problem of the beaming is the transmission. It is difficult to get high transmission and good beaming simultaneously. In our design, a single structure gives us good beaming and the Fabry-Pérot interference between two structures guarantees the total transmission. The two problems decouple. We can solve them separately. The two-layer structure also has the advantage in theoretical analysis. Comparing with the beaming of the photonic crystal channel, the two-layer structure gives similar results with much fewer rods. The complex channel modes are replaced by the simple Gaussian beam. The theoretical analysis will be much simpler.

In conclusion, we present a numerical analysis of the beaming and transmission of a Gaussian beam through a two-layer structure. This structure does not allow the Gaussian beam to diverge, and also gives a high transmission. By arranging several two-layer structures one can easily sustain the beaming and the high transmission for very large distances, of the order of 200λ. This simple design of beaming and transmission has advantage both in the theoretical analysis and in practical applications.

5.3 Control of beaming angles

5.3.1 Oblique beaming

Beside the research on forward beaming mentioned above, oblique beaming was also demonstrated using either metallic structures (75; 76) or photonic crystals (77). Control of the beaming angle is important in beaming research, which will give us much more flexibility in applications. I will present an efficient method to design a metallic structure which will steer the emitted field to any given angle here. I find the best beaming actually happens not at the forward direction but an oblique direction.

5.3.2 Introduction of a metallic beaming structure

Fig. 5.9 shows the structure studied here: a metallic subwavelength slit surrounded by grooves. It is an important plasmonic structure (78). As a two-dimensional structure, it can be solved analytically and serves as a theoretical model to understand the physics of beaming (53). It is a multi-function beaming structure. It can achieve both intermediate-field beaming (55) and far-field beaming (53) , specially far-field oblique beaming which I study here. The near-field focusing was also implemented by similar structures (79). Adding grooves along input surface will increase transmission dramatically (65; 80), so this structure can improve transmission and directionality of the output field simultaneously. Recently this structure was integrated with a semiconductor laser to converge the emitted fields (81).

Surface impedance boundary condition is used to replace the metal surface. Under TE polarization, the SIBC along $x = 0$ in Fig. 5.9 becomes

$$E_y + ZH_z = 0.$$

Define $F = E_y + ZH_z$. From now on in this section, the word "field " means $E_y + ZH_z$. We can define the far-field angular transmission $P(\theta)$ by the equation

$$P(\theta) = \lim_{r\to\infty} r|\vec{S}(\vec{r})|.$$

where $\vec{S}(\vec{r})$ is the Poynting vector at the point $\vec{r} = (r\sin\theta, r\cos\theta)$. Given the field along $x = 0$, $P(\theta)$ can be calculated as

$$P(\theta) = \frac{k_0}{2Z_0}\left|\frac{\cos\theta}{\cos\theta + Z/Z_0}\right|^2 \int_{-\infty}^{\infty} dy' F(0, y') e^{-ik_0 \sin\theta y'}. \tag{5.2}$$

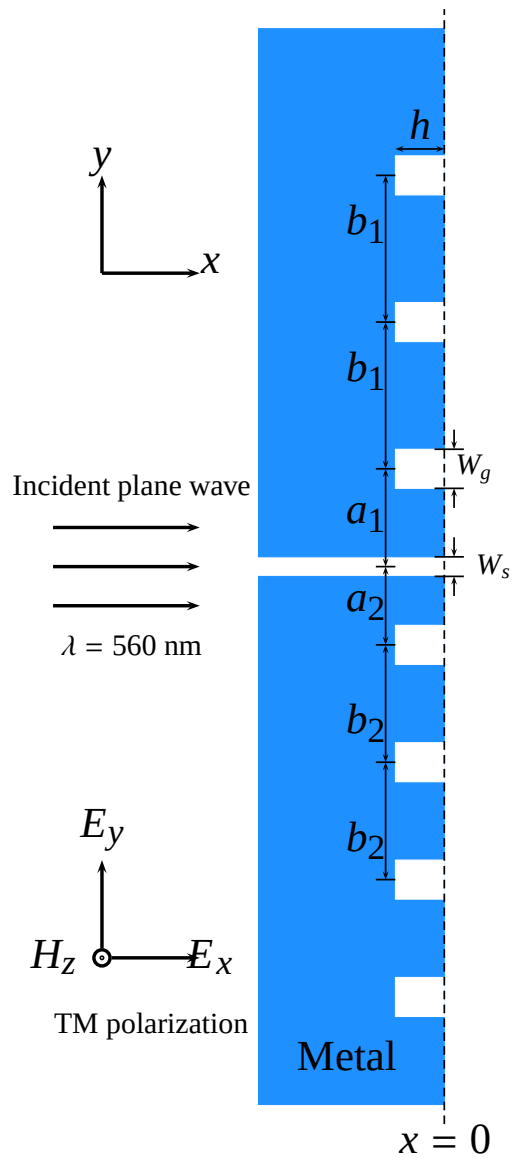

Figure 5.9 Schematic representation of the subwavelength metal slit surrounded by upper and lower periodic grooves.

where k_0 and Z_0 are the wave vector and the impedance in free space. Then the total transmission T is

$$T = \int_{-\pi/2}^{\pi/2} P(\theta) d\theta;$$

and the normalized far-field angular transmission $I(\theta) = P(\theta)/T$ is used to describe the far-field distribution.

Because of the surface impedance boundary condition, the filed along $x = 0$ is zero except the opening area of the slit and grooves. Since the widths of of the indentations (slit and grooves) are much smaller than the incident wavelength, we assume the field is a constant inside an indentation and Eq. (5.2) is basically the summation of several point sources. We notate the field at the ith grooves as F_i; positive i means the groove above the slit, negative i means the groove below the slit and $i = 0$ represents the slit.

By changing the geometry parameters of the grooves, we can change the field inside the grooves and control the output field distribution furthermore. It is impossible to exhaust all possible choices of grooves. We focus on the periodic grooves here. As Fig. 5.9 shows, all the grooves have the same shape with width $W_g = 40$ nm and depth $h = 100$ nm. The slit width is $W_s = 40$ nm. All the grooves above the slit consist in the upper periodic grooves described by the initial location a_1 and the period d_1. The grooves below the slit are also periodic and the corresponding parameters are a_2 and b_2. The grooves are symmetric when $a_1 = a_2$, $b_1 = b_2$ and we will use a and b to replace a_1, a_2 and b_1, b_2. In this paper we fix the wavelength $\lambda = 560$ nm and and try to find the best far-field oblique beaming at any given angle by modifying a_1, b_1, a_2 and b_2. Under the restriction of periodic grooves, we find an efficient algorithm to decide the parameters with explicit physical explanation.

5.3.3 Control of beaming angles

We begin our research by studying the fields at the opening area of the indentations when the grooves are symmetric. Figures 5.10(a) and 5.10(b) show the field's amplitude and phase when $a = b = 500$ nm. All the fields are normalized to the field at the slit exit, so $F_0 = 1$. We can find the field at the slit is much stronger than the fields at the grooves. The fields at the grooves decay smoothly with the distance to the slit and become very weak in the end. There are 56 grooves in each side of the

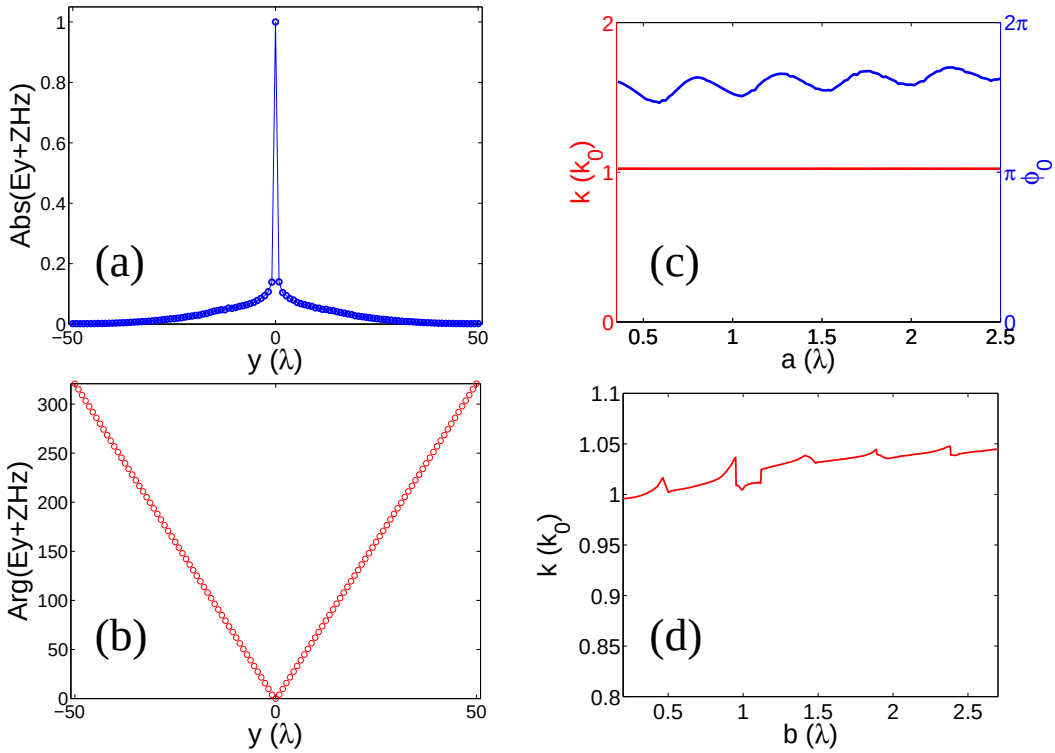

Figure 5.10 (a)&(b) Amplitudes and phases of fields at symmetric grooves with $a = b = 500$ nm; (c)&(d) the linear fitting results of phases of the fields at grooves for different a when $b = 500$ nm and for different b when $a = 500$ nm.

slit in this simulation. Further simulations show adding more grooves will not influence the far-field distribution much.

The most important property about the fields is the perfect straight line formed by the phases at the upper and lower periodic grooves in Fig. 5.10(b). So we can use a linear function to fit the phases at the upper half grooves. For a positive integer i, the location of the ith groove is $y_i = a + (i-1)b$ and the phase is $\Phi_i = \mathrm{Arg}(F_i)$, then

$$\Phi_i = ky_i + \phi_0. \tag{5.3}$$

This linear relation can be explained by the single scattering approximation. The surface plasmons are excited at the exit of the slit and propagate along the output surface. Then they will interact with the groove they meet. The weak fields at grooves suggests the interaction is weak. Most of the energy carried by the surface plasmons transmits through the grooves. It explains why the fields at grooves decay smoothly. Some energy is scattered into the free space and the rest is is reflected. The single scattering approximation neglects the reflected surface plasmons. The surface plasmons repeat their propagation and interaction process, so the phase change between two neighboring grooves is constant.

The single scattering approximation suggests that the fitting results will not change with a. Our simulations verified it. The fitting results are shown in Fig. 5.10(c) when $b = 500$ nm and a changes 200 nm to 1500 nm . We can see clearly that k is constant and the oscillation of ϕ_0 is weak.

Next we fix $a = 500$ nm and change b from 100 nm to 1500 nm. The slopes of the phases are shown is Fig. 5.10(d). The wavevector of the surface plasmons is $k_{sp} = 1.07k_0$. k_0 is the wave vector of the incident wave in free space. We can find k is close to but not equals k_{sp}. It is because of the small phase change introduced by the grooves. When b increases, which means less grooves along the surface, k converges to k_s.

The slope curve in Fig. 5.10(d) is not smooth when $kb \approx \pi n$ (n is an integer), which is because of "collective surface modes" (53). When $kb = \pi n$, the reflected surface plasmons from the all the grooves are in phase, so the single scattering approximation becomes weak. However, the single scattering approximation still works even at the collective surface modes since k is only a little bit away .

Now we have understand the phase difference between two neighboring grooves is kb. If

$$kb - k_0 b \sin\theta = 2\pi n, \tag{5.4}$$

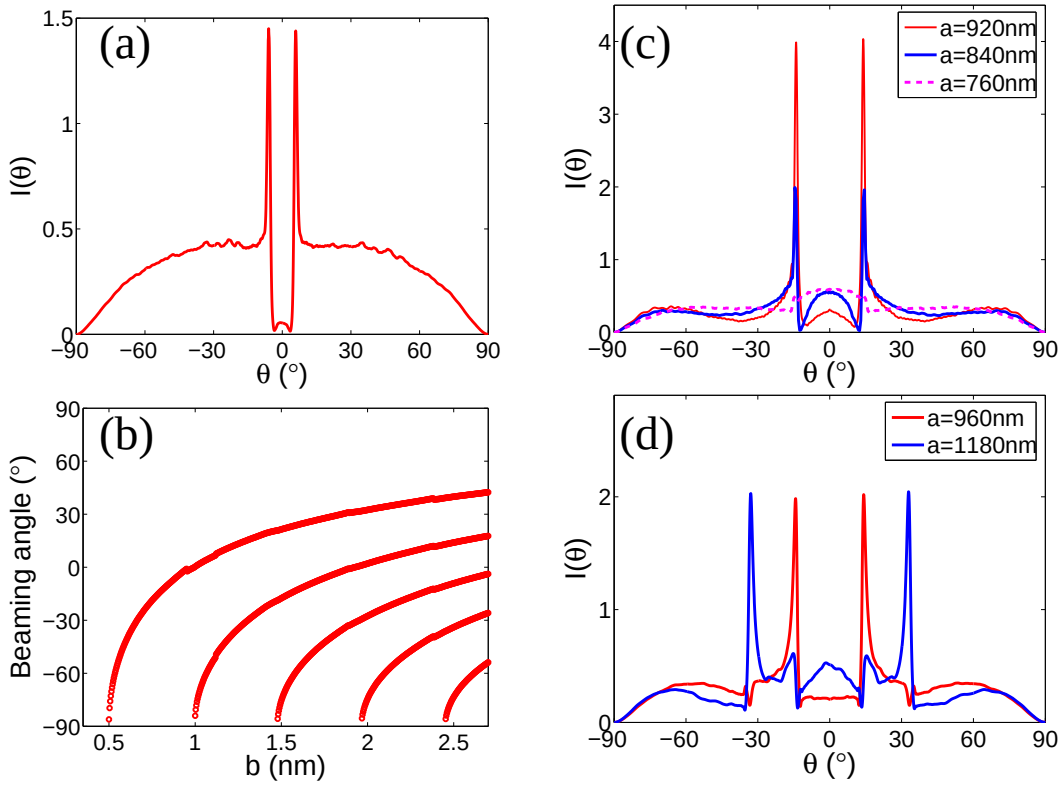

Figure 5.11 (a) Normalized far-field angular transmission when $a = b = 500$ nm; (b) Constructive interference angles at different groove periods; (c)&(d) Normalized far-field angular transmission when $b = 446$ nm and $b = 710$ nm.

and n is an integer, all the grooves above the slit have constructive far-field interference at the angle θ. Figure 5.4(b) shows the constructive interference angle of the upper periodic grooves with different period calculated from Equation (5.4).

We also calculate $I(\theta)$ from the fields at the indentations and the results are shown in Fig. 5.11(a). The normalized far-field angular transmission distribution curve has two peaks at $\pm 6°$, which are the constructive interference angles of the upper and lower half grooves.

Besides the two peaks, the transmission curve also has a smooth base. This is the contribution from the slit. The existence of the base means the energy is diffracted to all the angles, which is against our purpose of good directional emission. But it also give us the opportunity to control the peak heights, as shown in Fig. 5.11(c) and (d). Considering the contributions from the slit and the upper periodic grooves to the infinite at angle θ, the phase difference between them is $\phi_0 + ka - k_0 a \sin\theta$. So we can

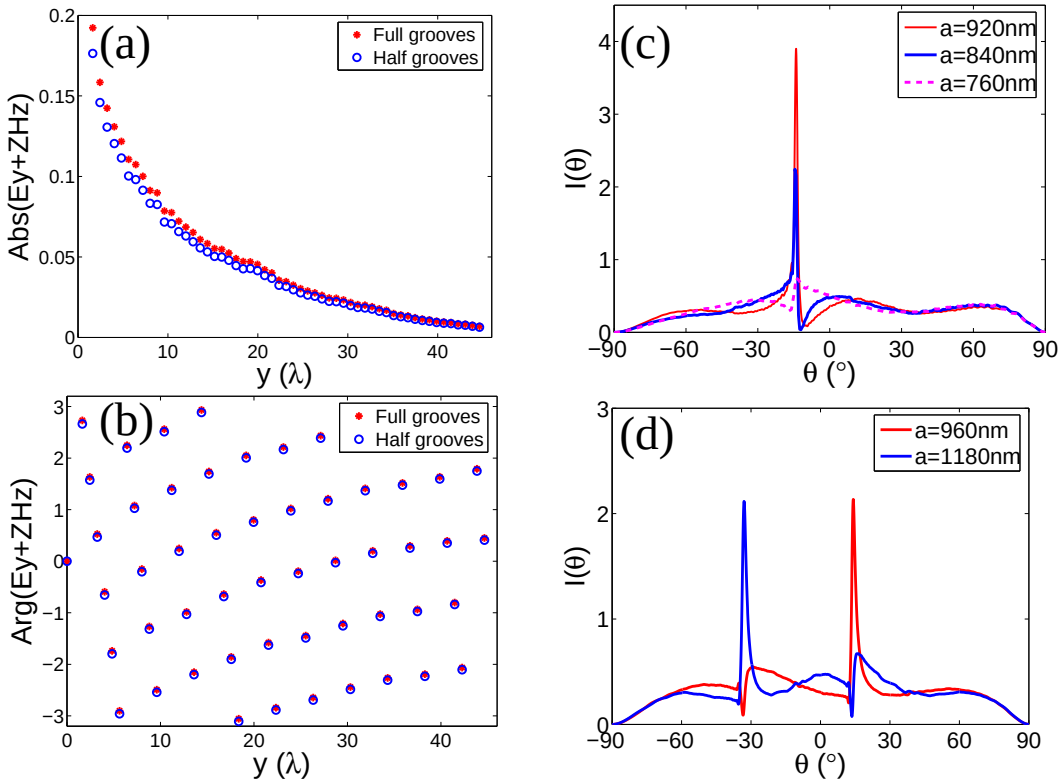

Figure 5.12 (a)&(b) Amplitudes and phases of fields at grooves above the slit when the grooves are symmetric with $a_1 = a_2 = 920$ nm, $b_1 = b_2 = 446$ nm and when there are only upper periodic grooves with $a_1 = 920$ nm, $b_1 = 446$ nm; (c)&(d) Normalized far-field angular transmission generate by the slit with upper periodic grooves alone when $b_1 = 446$ nm and $b_1 = 710$ nm.

control the interference and the peak heights by changing a.

Fig. 5.11(c) also shows the beaming angle is independent from a. It is only decided by the constructive interference angle of the upper and lower periodic grooves. In Fig. 5.11(d), $b = 710$ nm and there are two constructive interference angles. The value of a decides which angle becomes the beaming angle.

Until now we always use symmetric grooves, so the far-field distributions are also symmetric and have two peaks. The ideal oblique beaming has only one peak. A nature idea is to remove all the grooves on the one side of the slit and use only half grooves. Our simulations show this simple idea works.

Fig. 5.12(a) and (b) show the amplitude and phase of the fields at grooves above the slit is accompa-

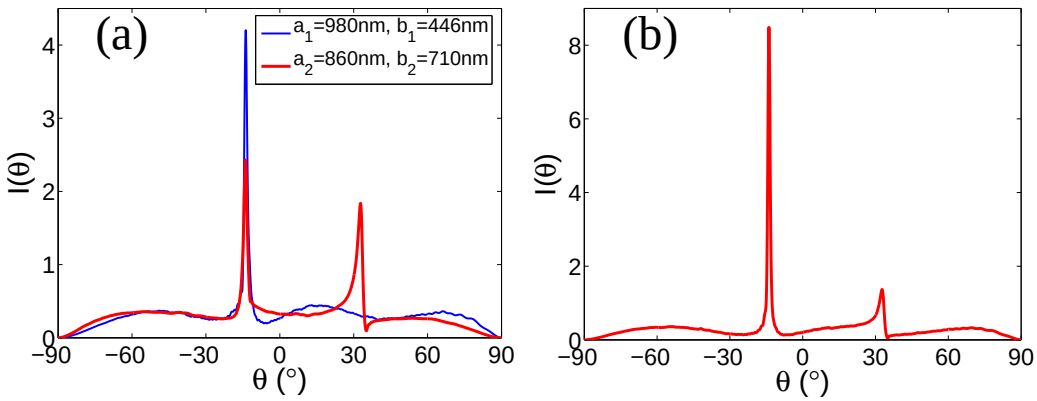

Figure 5.13 (a) Normalized far-field angular transmission of a slit with upper periodic grooves alone when $a_1 = 980$ nm, $b_1 = 446$ nm and a slit with lower periodic grooves alone with $a_2 = 860$ nm, $b_2 = 710$ nm; (b) Normalized far-field angular transmission of a slit with asymmetric upper and lower periodic grooves when $a_1 = 980$ nm, $b_1 = 446$ nm, $a_2 = 860$ nm, $b_2 = 710$ nm.

nied by upper grooves only or by symmetric upper and lower grooves. The fields under the two groove setting are very close to each other, specially the phase. The similarity suggests the independence between the upper and lower periodic grooves. This independence is also a nature results of single scattering approximation. Since we neglect the reflected surface plasmons, the upward plasmons have no chance to reach the grooves below the slit, and *vice versa*. Fig. 5.12(c) and (d) show the far-field distribution of upper grooves alone when $b_1 = 446$ nm and $b_{=}710$ nm. Comparing with Fig. 5.11(c) and (d), we can find that removing the lower grooves will not influence the peaks generated by the upper periodic grooves: the location, the height and and of the peaks keeps.

Fig. 5.11(b) shows the constructive interference angle changes from 0° to 90° when b_1 changes form 540 nm to 280 nm. Given any angle, the design of the beaming structure of a slit with half grooves has two steps: first we check Fig. 5.11(b) to get the period of grooves b_1 smaller than 540 nm; then we fix the period and scan the initial location of the grooves a_1 to maximize the transmission peak height.

Fig. 5.11(b) also shows when $b_1 > 540$ nm, the upper periodic grooves have several constructive interference angles and some of them are positive. It also means the lower periodic grooves with period $b_2 > 540$ nm can have negative constructive interference angles. The upper periodic grooves with

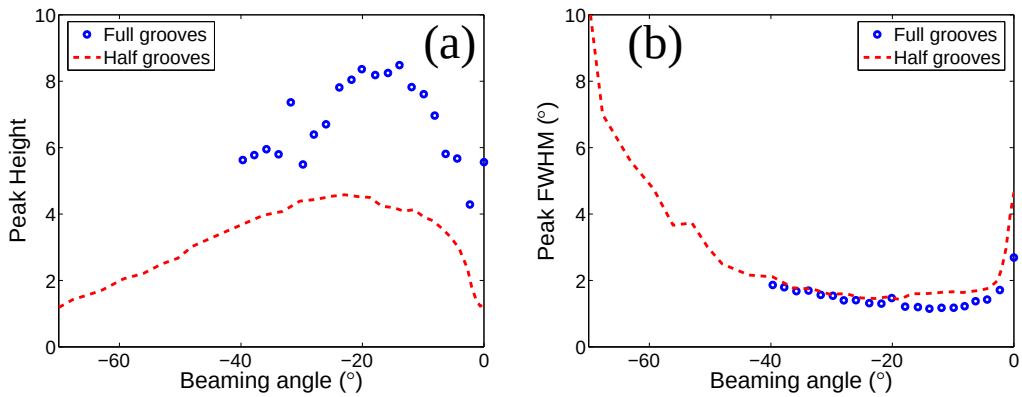

Figure 5.14 (a) Heights and (b) full-widths at half-maximum of the highest far–field transmission peak at different angles generated by upper periodic grooves only and asymmetric upper and lower periodic grooves.

$b_1 = 446$ nm and the lower periodic grooves with $b_2 = 710$ nm both have the constructive interference angle at $-14°$. We can choose appropriate initial locations to maximize the peaks at the angle and the far-field distributions of the half grooves are shown in Fig. 5.13(a). Then we put the upper and lower periodic grooves together and the far-field distribution of the asymmetric full grooves is shown in Fig. 5.13(b). Now all the three components, the upper periodic grooves, the lower periodic grooves and the slit, interfere constructively. So we get a very strong far-field transmission peak at $\theta = -14°$.

Fig. 5.14 shows the properties of the highest transmission peaks we can achieve at different beaming angle using upper periodic grooves only or using asymmetric upper and lower periodic grooves. Adding grooves below the slit improves the transmission peaks. The highest peaks are achieved round 20°. When the beaming angle is small, the surface waves enter the collective surface modes; when the beaming angle is big, $\cos\theta$ in the numerator of Eq. (5.2) suppresses the peaks.

5.3.4 Frequency splitter

We can also design a frequency splitter using asymmetric grooves: two wavelengths enter the slit and leave at different angles. We have understood the upper periodic groove has only one possible beaming angle when its period is smaller than the wavelength and the angle satisfies

$$kb - k_0 b \sin\theta = 2\pi. \tag{5.5}$$

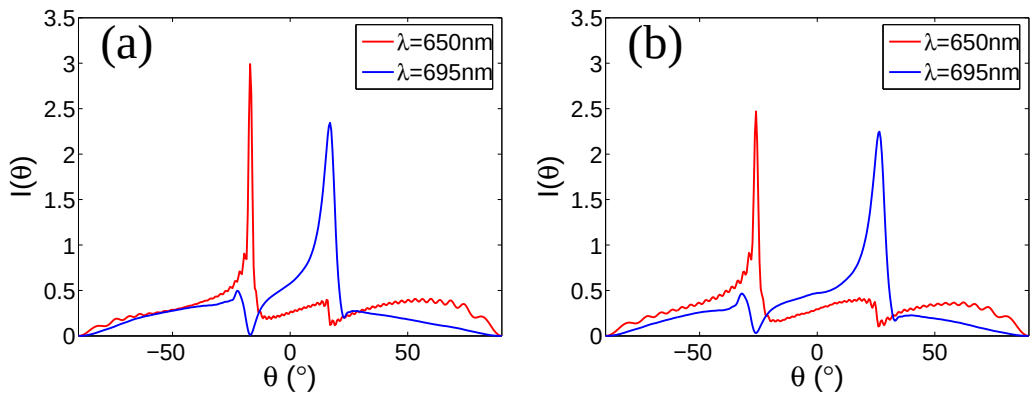

Figure 5.15 Normalized far-field angular transmission at two frequencies when (a) $a_1 = 600$ nm, $b_1 = 500$ nm, $a_2 = 400$ nm, $b_2 = 500$ nm and (b) $a_1 = 540$ nm, $b_1 = 450$ nm, $a_2 = 360$ nm, $b_2 = 450$ nm. The groove depth is $h = 500$ nm.

The phase difference between the upper periodic grooves and the slit at the beaming angle is

$$ka + \phi_0 - k_0 a \sin\theta = 2\pi a/b + \phi_0. \tag{5.6}$$

We can find two frequencies λ_1 and λ_2 satisfying $\phi(\lambda_1) - \phi(\lambda_2) = \pi$. When the upper grooves and slit have constructive interference and result in a enhanced transmission peak at one wavelength, the grooves and the slit have destructive interference and suppress the transmission peak at the other wavelength.

In Fig. 5.15(a), the two frequencies are 650 nm and 695 nm. When the depth of grooves is $h = 100$ nm, the difference between ϕ_0 at the two wavelengths is 1.7. To get bigger difference, we have to increase h to 500 nm. The period of the upper periodic grooves is $b_1 = 500$ nm. The constructive interference angles of the upper grooves are around $-17°$ for both frequencies. But because of the interference between the grooves and slit, only the field with $\lambda = 650$ nm has a transmission peak at its constructive interference angle. $b_2 = b_1$ in Fig. 5.15(a), so the constructive interference angles of the lower periodic grooves are around 17°. We choose a different a_2 to enhance the peak of the field with $\lambda = 695$ nm. Though the two incident wavelengths are quite close to each other, the angular transmission peaks are 34° away from each other.

Based on Eq. (5.5) and (5.6), if we change b but keep a/b unchanged, the locations of the peaks changes but the phase difference between the grooves and the slit keeps. In Fig. 5.15(b), a_1, a_2, b_1

and b_2 are 10% smaller then the corresponding values in Fig. 5.15(a). The structure still works as a frequency splitter and the output peaks are 52° away.

5.3.5 Conclusion

I have shown that a metallic slit surrounding by grooves is a good oblique beaming structure. I implement good oblique beaming at any angle in the range of ±70° using a metallic subwavelength slit with one-side periodic grooves and better oblique beaming at an angle between ±40° using a slit surrounded by upper and lower periodic grooves. I also design a frequency splitter based on the structure.

APPENDIX A. Simulations of infinitely long waveguides

A.1 Perfect matched layer

Computers can not simulate an area with finite size. But sometimes we need simulate the interaction between electromagnetic waves and objects located in an unbounded region. For these problems, we have to use absorbing boundary conditions (ABC) to terminate the simulation area (82).

Perfect matched layer (PML) is one kind of absorbing boundary conditions. The basic idea is to use a lossy medium surrounding the simulation area and ideally the lossy medium will absorb incident waves without any reflections regardless the incident angles, frequnencies and polarizations. Comparing with other analytic absrorbing boundary conditions, PML ABS uses a litter bit larger simulation area to contain the lossy medium but generally reflects less energy back. Normally the thickless of PML is about several wavelengths.

Below I will introduces an implementation of PML ABC: a uniaxial PML (UPML), which mechanics can be proved analytically from Maxwell'e equations (83). Fig. A.1 shows the schematic of the problem: the $x < 0$ half-plane is the unbounded simulation area and contains isotropic dielectric; the $x > 0$ half-plane is the perfect matched layer. We will study the wave propagation from the dielectric to the UPML. The dielectirc has scalar permittivity ϵ_1 and permeability μ_1. The UPML is comprised of a uniaxial anistropic medium and has the permittivity and permeability tensor

$$\bar{\bar{\epsilon}}_2 = \epsilon_1 \bar{\bar{s}}; \quad \bar{\bar{\mu}}_2 = \mu_1 \bar{\bar{s}}; \quad \bar{\bar{s}} = \begin{bmatrix} s_x^{-1} & 0 & 0 \\ 0 & s_x & 0 \\ 0 & 0 & s_x \end{bmatrix}; \quad s_x = 1 + i\sigma_x/\omega\epsilon_1. \tag{A.1}$$

Suppose a plane waves in the UPML has magnetic fields as

$$\vec{H}_2 = \vec{H}_0 \exp\left[i(k_{2x}x + k_{2y}y - \omega t)\right]. \tag{A.2}$$

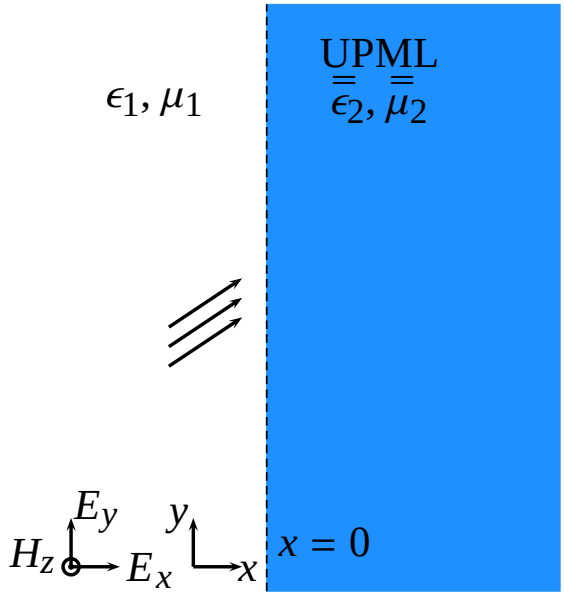

Figure A.1 Schematic representation of incidence from isotropic dielectric in $x < 0$ half-plane to perfect matched layer in $x > 0$ half-plane.

Then the electric fields are

$$\vec{E}_2 = \vec{E}_0 \exp\left[i(k_{2x}x + k_{2y}y - \omega t)\right]. \tag{A.3}$$

and $\vec{E}_0$ and $\vec{H}_0$ satisfy

$$\vec{k}_2 \times \vec{E}_2 = \omega\bar{\bar{\mu}}_2\vec{H}_0; \qquad \vec{k}_2 \times \vec{H}_2 = -\omega\bar{\bar{\epsilon}}_2\vec{E}_0. \tag{A.4}$$

$$\vec{k}_2 = k_{2x}\hat{x} + k_{2y}\hat{y}. \tag{A.5}$$

I use TE polarization as a example here, so H_z, E_x and E_y are the only three non-zero components of electromagnetic waves. The equations about TM polarization are very similar.

Then the magnetic fields of the plane wave in the UPML under TE polarization become

$$\vec{H}_0 = H_0\hat{z}. \tag{A.6}$$

Based on the second equation of Eq. (A.4), we have

$$\vec{E}_0 = \left(-\frac{k_{2y}s_x}{\omega\epsilon_1}H_0, \quad \frac{k_{2x}}{\omega\epsilon_1 s_x}H_0, \quad 0\right). \tag{A.7}$$

Put the expression of $\vec{E}_0$ back to the first equation of Eq. (A.4), we get the dispersion relation of the plane waves in the UPML:

$$k_{2x}^2 = s_x^2(\omega^2\epsilon_1\mu_1 - k_{2y}^2). \tag{A.8}$$

The plane wave in the UPML is excited by the incident waves from dielectric side. Suppose the incident wave in the dielectric is

$$\vec{H}^{inc} = \hat{z}H_i \exp[i(k_{1x}x + k_{1y}y - \omega t)]. \tag{A.9}$$

Then the reflected wave is

$$\vec{H}^{ref} = \hat{z}RH_i \exp[i(-k_{1x}x + k_{1y}y - \omega t)]. \tag{A.10}$$

(R is the reflection coefficient) and the total fields in the dielectric are

$$\vec{H}_1 = \vec{H}^{inc} + \vec{H}^{ref} = \hat{z}H_i \exp[i(k_{1x}x + k_{1y}y - \omega t)] + \hat{z}RH_i \exp[i(-k_{1x}x + k_{1y}y - \omega t)]. \tag{A.11}$$

The electric fields in the dielectric are

$$\vec{E}_{1x} = -\frac{k_{1y}}{\epsilon_1\omega}\Big(H_i \exp[i(k_{1x}x + k_{1y}y - \omega t)] + RH_i \exp[i(-k_{1x}x + k_{1y}y - \omega t)]\Big); \tag{A.12}$$

$$\vec{E}_{1y} = \frac{k_{1x}}{\epsilon_1\omega}\Big(H_i \exp[i(k_{1x}x + k_{1y}y - \omega t)] - RH_i \exp[i(-k_{1x}x + k_{1y}y - \omega t)]\Big); \tag{A.13}$$

$$\vec{E}_{1z} = 0. \tag{A.14}$$

H_z and E_y are continuous along the interface $x = 0$. So $k_{2y} = k_{1y}$ and k_{2x} is decided by the dispersion relation of Eq. (A.8). Then the fields along $x = 0$ are

$$H_{1z}(0, y) = (1 + R)H_i \exp[i(k_{1y}y)]; \tag{A.15}$$

$$H_{2z}(0, y) = H_0 \exp[i(k_{1y}y)]; \tag{A.16}$$

$$E_{1y}(0, y) = \frac{k_{1x}}{\epsilon_1\omega}(1 - R)H_i \exp[i(k_{1y}y)]; \tag{A.17}$$

$$E_{2y}(0, y) = \frac{k_{2x}}{\epsilon_1\omega s_x}H_0 \exp[i(k_{1y}y)]. \tag{A.18}$$

and the dispersion relations are

$$k_{1x}^2 = \omega^2\epsilon_1\mu_1 - k_{1y}^2; \qquad k_{2x}^2 = s_x^2(\omega^2\epsilon_1\mu_1 - k_{1y}^2). \tag{A.19}$$

From $H_{1z}(0, y) = H_{2z}(0, y)$, we get $(1 + R)H_i = H_0$. From $E_{1z}(0, y) = E_{2z}(0, y)$, we get $(1 - R)^2H_i^2 = H_0^2$. Obviously $R = 0$ is the only choice to make the two equations correct and when $R = 0$, $H_0 = H_i$, $k_{2x} = k_{1x}s_x$, the continuities of H_z and E_y along $x = 0$ are correct for any frequency and incident

angle. So we proved the UPML perfectly matches the dielectric and generates no reflection for any frequencies and incident angles.

When incident angle is θ, we have

$$k_{1x} = k\cos\theta, \quad k_{1y} = k\sin\theta, \quad k = \omega\sqrt{\epsilon_1\mu_1}, \quad \theta \in (-\pi/2, \pi/2). \tag{A.20}$$

The incident wave is

$$\vec{H}_1 = \hat{z}H_i \exp[i(k\cos\theta x + k\sin\theta y - \omega t)]; \tag{A.21}$$

$$\vec{E}_1 = \left(-\hat{x}Z_d\sin\theta + \hat{y}Z_d\cos\theta\right)H_i \exp[i(k\cos\theta x + k\sin\theta y - \omega t)]; \tag{A.22}$$

$$Z_d = \sqrt{\mu_1/\epsilon_1}. \tag{A.23}$$

and the fields in the UPML is

$$\vec{H}_2 = \hat{z}H_i \exp[i(k\cos\theta x + k\sin\theta y - \omega t)]\exp[-\sigma_x Z_d\cos\theta x]; \tag{A.24}$$

$$\vec{E}_2 = \left(-\hat{x}s_x Z_d\sin\theta + \hat{y}Z_d\cos\theta\right)H_i \exp[i(k\cos\theta x + k\sin\theta y - \omega t)]\exp[-\sigma_x Z_d\cos\theta x]. \tag{A.25}$$

When $\sigma_x > 0$, the UPML is lossy and the fields decay exponentially inside. The perfect matching property and the lossy property make UMPL a ideal choice to truncate the unbounded simulation area.

A.2 Perfect electric conductor waveguide

In the previous section we have introduced UPML for unbounded simulation area. Now we begin to study waveguides with infinite length, which is partially unbounded. The diagram is shown in Fig. A.2. The two-dimensional waveguide is bounded by two PEC surfaces along y direction and the waveguide width is W. Now I will prove inserting UPML in the waveguide will make the waveguide infinite along x direction.

The fields in a waveguide are the summation of the eigenmodes. Suppose $\vec{H}_1$ is a eigenmode in the waveguide propagating at $+x$ direction. I consider only TE mode here. So

$$\vec{H}_1 = \hat{z}(A\exp[ik_y y] + B\exp[-ik_y y])\exp[i(k_x x - \omega t)]. \tag{A.26}$$

The eigonmode in the waveguide is always the summation of two plane waves. Based on the knowledege of UPML (Eq. (A.21) to Eq. (A.25)) in the previous section, we know the field distributions

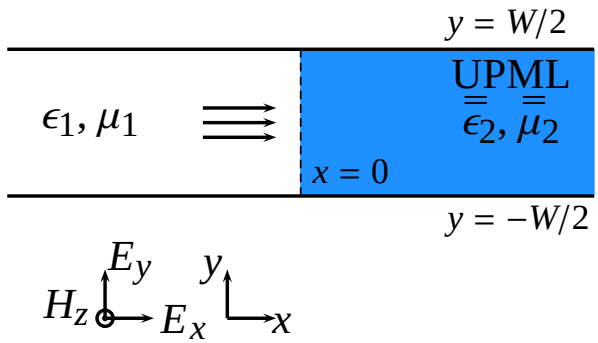

Figure A.2 Schematic representation of a UPML in a PEC or a SIBC waveguide.

below ($\vec{H}_1$ and $\vec{H}_1$ are the fields in the dielectric; $\vec{H}_2$ and $\vec{H}_2$ are the fields in the dielectric;) satisfy the boundary condition along $x = 0$, which is the interface between the dielectric and UPML.

$$\vec{H}_1 = A\hat{z}\exp[i(k_x x + k_y y - \omega t)] + B\hat{z}\exp[i(k_x x - k_y y - \omega t)]; \tag{A.27}$$

$$\begin{aligned}\vec{E}_1 = {} & A\Big(-\hat{x}Z_d k_y/k + \hat{y}Z_d k_x/k\Big)\exp[i(k_x x + k_y y - \omega t)] \\ & + B\Big(\hat{x}Z_d k_y/k + \hat{y}Z_d k_x/k\Big)\exp[i(k_x x - k_y y - \omega t)];\end{aligned} \tag{A.28}$$

$$\begin{aligned}\vec{H}_2 = {} & A\hat{z}\exp[i(k_x x + k_y y - \omega t)]\exp[-\sigma_x Z_d k_x x/k] \\ & + B\exp[i(k_x x - k_y y - \omega t)]\exp[-\sigma_x Z_d k_x x/k];\end{aligned} \tag{A.29}$$

$$\begin{aligned}\vec{E}_2 = {} & A\Big(-\hat{x}s_x Z_d k_y/k + \hat{y}Z_d k_x/k\Big)\exp[i(k_x x + k_y y - \omega t)]\exp[-\sigma_x Z_d k_x x/k] \\ & + B\Big(\hat{x}s_x Z_d k_y/k + \hat{y}Z_d k_x/k\Big)\exp[i(k_x x - k_y y - \omega t)]\exp[-\sigma_x Z_d k_x x/k];\end{aligned} \tag{A.30}$$

Here $k = \sqrt{k_x^2 + k_y^2}$.

At a PEC boundary, the electric fields are perpendicular to the boundary. So the eigenmode in the PEC waveguide satisfies

$$\begin{aligned}A\exp[ik_y W/2] - B\exp[-ik_y W/2] = 0; \\ A\exp[-ik_y W/2] - B\exp[ik_y W/2] = 0.\end{aligned} \tag{A.31}$$

Given boundary conditions Eq. (A.31), it is easy to verify the electric fields in UPML are also perpendicular to the boundaries. So Eq. (A.27) to Eq. (A.30) are a possible field distribution of UPML in the PEC waveguide if the incident wave is $\vec{H}_1$. Becuse of the uniqueness theorem of the field distribution, they are the only solution. Obviously there isn't any reflected wave in the solution and the

fields in UPML decay exponentially. So inserting UPML in a PEC waveguide makes the waveguide infinitely long.

A.3 Metallic waveguide

Now I study how to implement an infinitely-long metallic waveguide. The geometry is the same as the PEC waveguide shown in Fig. A.2. The difference is now the waveguide is bounded by surface impedance boundaries.

Using the same method shown in the previous section, I begin from an eigenmode in the waveguide and Eq. (A.27) to Eq. (A.27) are a field distribution satisfying the boundary condition along the dielectric-UPML interface.

Now let's consider the boundary conditions along the waveguide boundaries. SIBC requires the eigenmode satisfies

$$\frac{E_{1x}(x, W/2)}{H_{1z}(x, W/2)} = -Z_m \quad \text{and} \quad \frac{E_{1x}(x, -W/2)}{H_{1z}(x, -W/2)} = Z_m.$$

Here Z_m is the surface impedance. So I get$A$ and B satisfy

$$\begin{aligned}
\frac{A\exp[ik_yW/2] - B\exp[-ik_yW/2]}{A\exp[ik_yW/2] + B\exp[-ik_yW/2]} &= \frac{Z_m k}{Z_d k_y};\\
\frac{A\exp[-ik_yW/2] - B\exp[ik_yW/2]}{A\exp[-ik_yW/2] + B\exp[ik_yW/2]} &= -\frac{Z_m k}{Z_d k_y}.
\end{aligned}$$

Then the fields in the UPML at the boundaries satisfy

$$\begin{aligned}
\frac{E_{2x}(x, W/2)}{H_{2z}(x, W/2)} &= \frac{-As_xZ_dk_y/k\exp[ik_yW/2] + Bs_xZ_dk_y/k\exp[-ik_yW/2]}{A\exp[ik_yW/2] + B\exp[-ik_yw/2]}\\
&= -\frac{s_xZ_dk_y}{k}\frac{A\exp[ik_yW/2] - B\exp[-ik_yw/2]}{A\exp[ik_yW/2] + B\exp[-ik_yw/2]} = -s_xZ_m;\\
\frac{E_{2x}(x, -W/2)}{H_{2z}(x, -W/2)} &= \frac{-As_xZ_dk_y/k\exp[-ik_yW/2] + Bs_xZ_dk_y/k\exp[ik_yw/2]}{A\exp[-ik_yW/2] + B\exp[ik_yw/2]}\\
&= -\frac{s_xZ_dk_y}{k}\frac{A\exp[-ik_yW/2] - B\exp[ik_yw/2]}{A\exp[-ik_yW/2] + B\exp[ik_yw/2]} = s_xZ_m;
\end{aligned}$$

So the fields in the UPML also satisfy SIBC at the waveguide boundaries but the surface impedance is s_xZ_m.

In conclusion, inserting UPML in a SIBC waveguide with surface impedance Z_m and changing the surface impedance of boundaries bounding UPML to s_xZ_m makes the waveguide infinitely long.

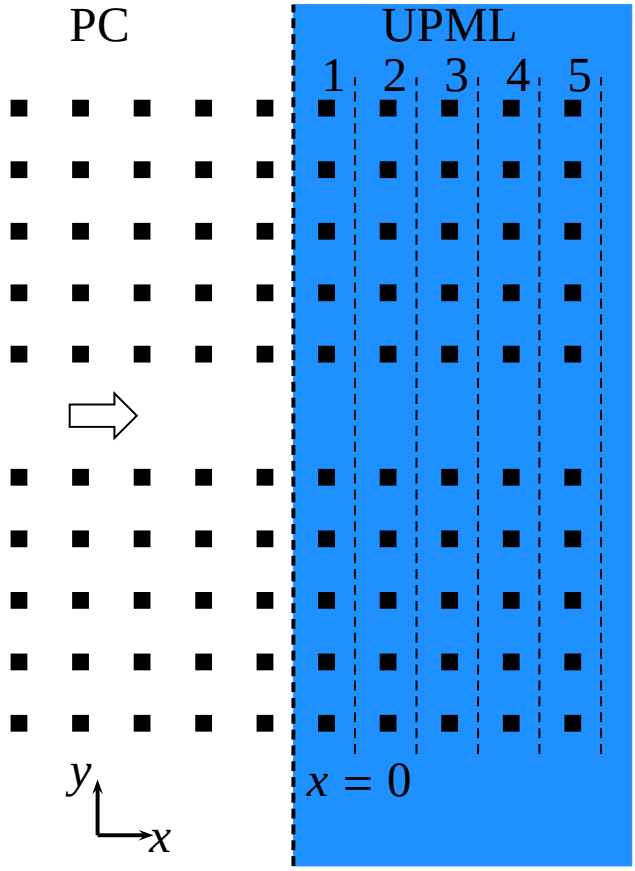

Figure A.3 Schematic representation of a two-dimensional photonic crystal waveguide with infinite length. The left region is an ordinary photonic crystal (PC) waveguide; the right blue region is the PC-based uniaxial perfectly matched layer (UPML). $x = 0$ is the PC-UPML interface.

When $Z_s = 0$, SIBC becomes PEC and the corresponding UPML boundaries is PEC too, which is the conclusion of the previous section.

A.4 Photonic waveguide

In previous section, I have implemented PEC or SIBC waveguides with infinite length by intserting UPML in the waveguide and proved the inserted UPML absorbs the incident wave without reflection analytically. UPML can also be used to simulated the infinitely-long photonic crystal waveguide (84; 85).

Fig. A.3 shows an example of a photonic crystal waveguide followed by a PC-based UPML. The

UPML begins from $x = 0$. The region on the left of the $x = 0$ line is an ordinary two-dimensional photonic crystal waveguide. In this example, the photonic crystal is composed by a suqare arrry of rods in air. A line defect is introduced into the photonic crystal to form the waveguide. The region on the right of the $x = 0$ line is the UPML. Instead of the homogeneous UPML used to terminate a PEC or SIBC waveguide, here the UPML keeps the geometry of the PC waveguide before it. The UPML in the rods have the permittivity and permeability tensor $\bar{\bar{\epsilon}}_{rod} = \epsilon_d \epsilon_0 \bar{\bar{s}}$ and $\bar{\bar{\mu}}_{rod} = \mu_0 \bar{\bar{s}}$; the UPML around the rods are $\bar{\bar{\epsilon}}_{air} = \epsilon_0 \bar{\bar{s}}$ and $\bar{\bar{\mu}}_{air} = \mu_0 \bar{\bar{s}}$. Here ϵ_d is the relative permittivity of the rods of the photonic crystal and

$$\bar{\bar{s}} = \begin{bmatrix} s_x^{-1} & 0 & 0 \\ 0 & s_x & 0 \\ 0 & 0 & s_x \end{bmatrix} \tag{A.32}$$

The UPML in the rods perfectly matches the rods of the ordinary PC and the UPML surronding the rods perfectly matches the air. This UPML is called PC-based UPML.

To suppress the reflection, s_x is a function of the location. s_x can change continuously. But for simplicity, it changes discreetly in my simulations. As shown in Fig. A.3, the UPML is cut into several layers and each layer is numbered form left to right. s_x keeps constant in one layer and changes with the layer number. Nomally the function is

$$s_x(n) = 1 + iM(n/N)^p.$$

Here n is the layer number; N is the total number of UPML layers; p is a postive integer.

Besides the infinitely-long photonic crystal, the PC-based UPML can be used to simulate any infinite PC structures with translation symmetry. In Fig. A.5 and Fig. A.6, I study the transmissio into a photonic crystal grating with infinite length, which is shown in Fig. A.4. The photonic crystal is a square array of square alumina rods. The lattice constant is $a = 11$ mm. The square rods are with side length $d = 3.1mm$ and relative permittivity $\epsilon_d = 9.8$. This photonic crystal has a band gap between 9.43 and 12.78 GHz under TE polization. There are periodic slits inside the photonic crystal every four columns of rods. The width of the slits is $W = 15$ mm. The grating is illuminated by normal incident plane wave with frequency $f = 9.6$ GHz, which is in the band gap. The total number of layes of the photonic crystal changes from 20 to 40. The total number of layers of the UPML is $N = 20$.

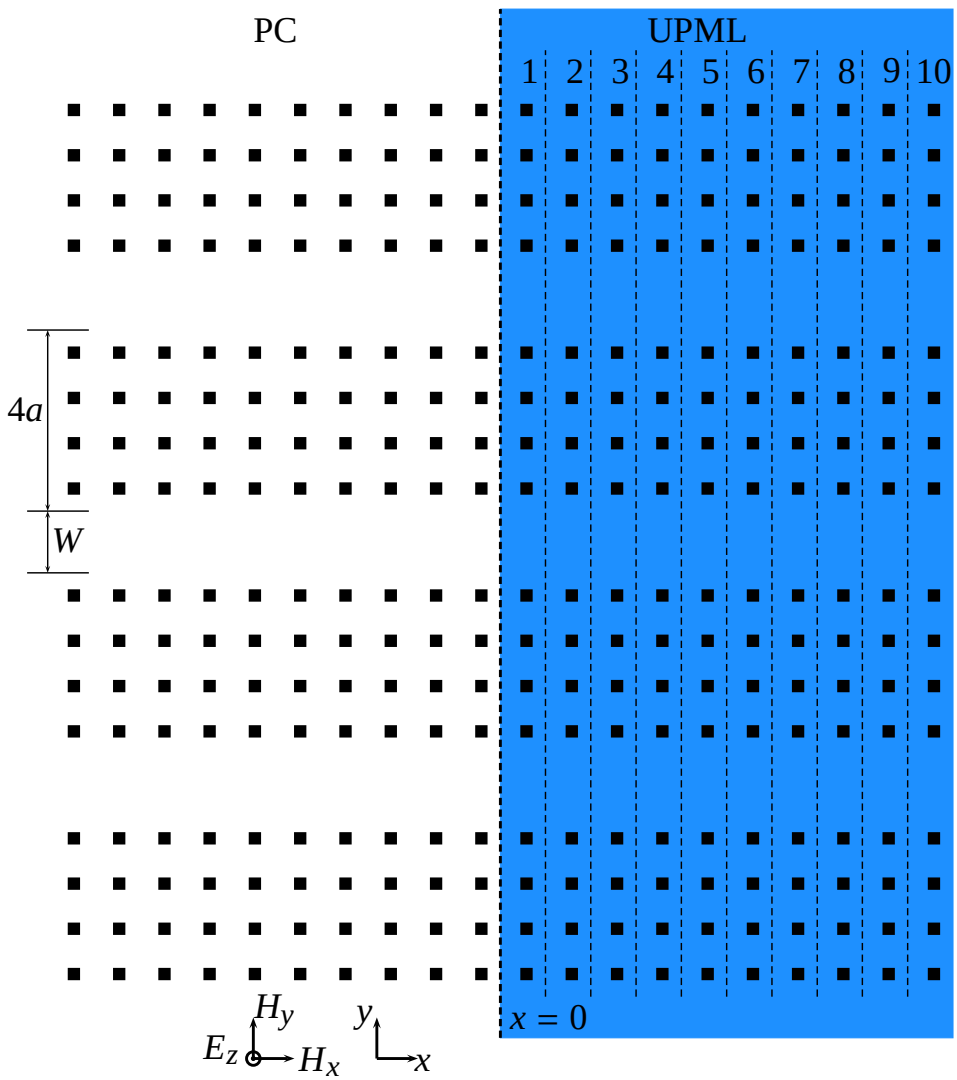

Figure A.4 Schematic representation of a two-dimensional photonic crystal grating with infinite length. The left region is the ordinary photonic crystal (PC) grating; the right blue region is the PC-based uniaxial perfectly matched layer (UPML). $x = 0$ is the PC-UPML interface. The incident wave is TM polarized.

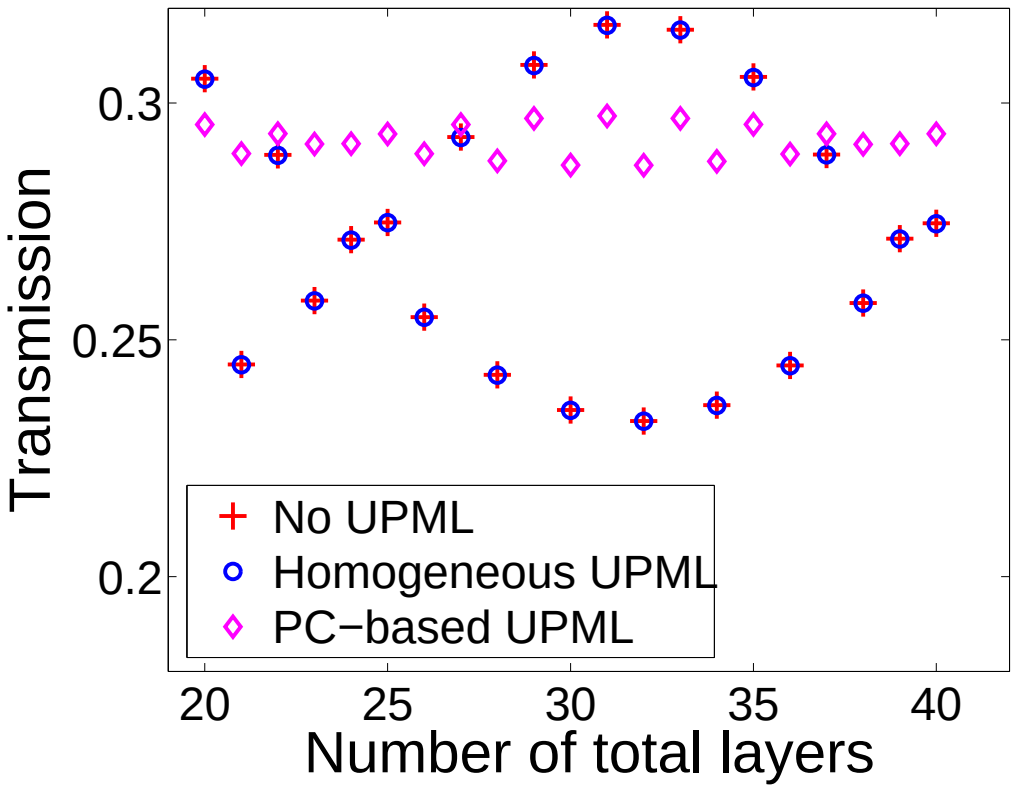

Figure A.5 Transmsission into a photonic crystal grating followed by free space, a homogeneous UPML and a PC-based UPML. The homogeneous UPML has $\bar{\bar{\epsilon}}_{rod} = \bar{\bar{\epsilon}}_{air} = \epsilon_0 \bar{\bar{s}}$ and $s_x = 1 + 0.3i$. The PC-based UPML has $N = 20, M = 0.3, p = 0$. The total number of layers of the graing changes from 20 to 40.

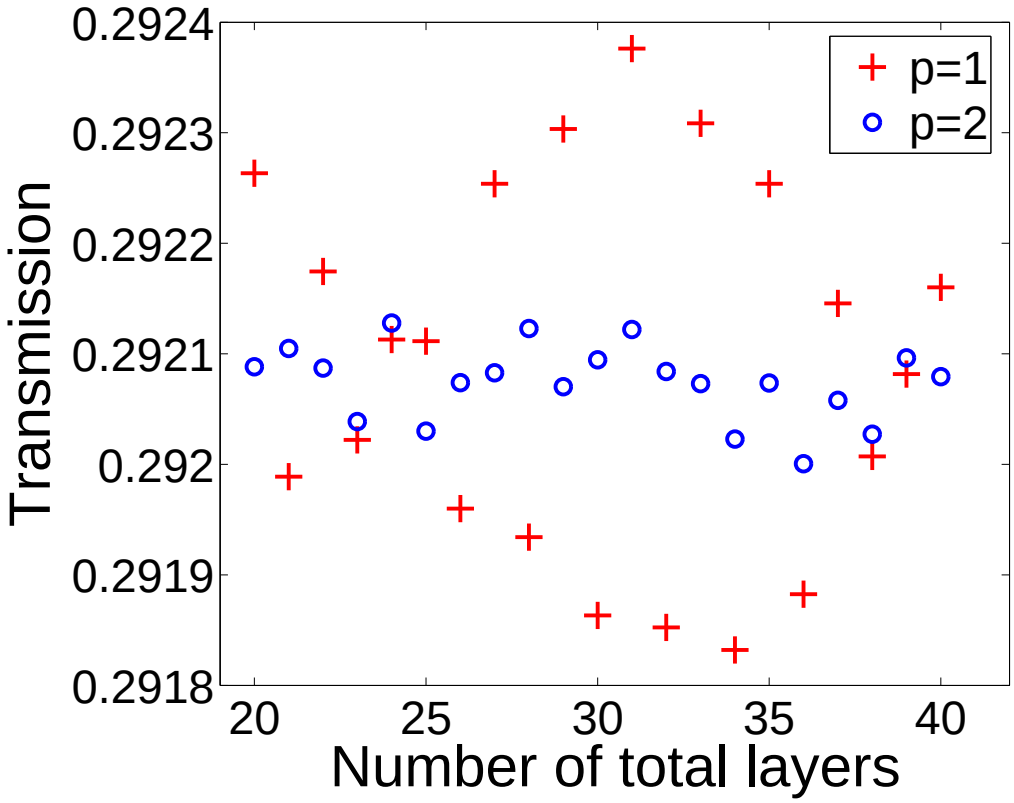

Figure A.6 Transmsission into a photonic crystal grating followed by PC-based UPMLs. The parameters of the PC-based UPMLs are $N = 20, M = 0.3, p = 1$ and $N = 20, M = 0.3, p = 2$. The total number of layers of the graing changes from 20 to 40.

In Fig. A.5, I show the transmission into the grating followed by air and homogeneous UPML. The homogeneous UPML means the permittivity tensor of the rods in UPML is the same as the tensor of the UPML around it; $\bar{\bar{\epsilon}}_{rod} = \bar{\bar{\epsilon}}_{air} = \epsilon_0 \bar{\bar{s}}$ and $s_x = 1 + 0.3i$. The two curves overlap with each other. It shows the homogeneous UPML is equivalent to free space. The transmission curve oscillates with the total number of layers of the ordinary photonic crystal. It is because of the Fabry-Pérot interference between the input and output surfaces of the photonic crystal grating. The oscillation amplitude descirbes the reflection from the output surface. Then I put a PC-based UPML with $M = 0.3, p = 0$ after the grating. The oscillation amplitude decreases dramatically, which means the UPML matches the grating very well. We should notice that this mathc is not perfect.

To get better simulation results, we have to consider two things; First the permittivity and permeability difference between two adjacent layers should be mall to suppress the reflection since the match isn't perfect; second, s_x in the UPML should be big to absort the incident wave efficiently. So in Fig. A.6, I increse s_x slowly. I put the PC-based UPML with $M = 0.3, p = 1$ and with $M = 0.3, p = 2$ after the grating. The oscillation amplitude becomes very small, specially when $p = 2$. So the grating followed by a PC-based UPML is equivalent to a grating with infinite length. We can continue to try differet function of s_x to get the optimal results.

BIBLIOGRAPHY

[1] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer-Verlag, 1988).

[2] S. A. Maier, *Plasmonics: Fundamentals and Applications* (Springer-Verlag, 2007).

[3] N. W. Ashcorft and N. D. Mermin, *Solid State Physics*, (Holt, Rinehart and Winston, 1976).

[4] M. P. Marder, *Condensed Matter Physics*, (HJohn Wiley & Sons, Inc., 2000).

[5] P.B. Johnson and R.W. Christy, "Optical Constants of the Noble Metals", Phys. Rev. B **6**, 4370 (1972).

[6] J. D. Joannopoulos, S. G. Johnson, J. N. Winn and R. D. Meade , *Photonic Crystal: Molding the Flow of Light* Second Edition (Princeton University Press, 2008).

[7] K. Sakoda, *Optical Properties of Photonic Crystal*, (KSpringer, 2004).

[8] J. Lourtioz, *Photonic Crystals: Towards Nanoscale Photonic Devices* Second Edition (Springer-Verlag, 2008)

[9] M. Born and E. Wolf, *Principles of Optics*, Seventh Edition, (Pergamon Press Ltd, London, 1999).

[10] C. M. Soukoulis, S. Lindedn and M. Wegener, "Negative refractive index at optical wavelengths", Science **315**, 47 (2007).

[11] N. Engheta and R. W. Ziolkowski (2006-06). *Metamaterials: physics and engineering explorations*, (John Wiley & Sons, Inc., 2006).

[12] E. N. Economou, "Surface plasmons in thin films", Phys. Rev. **182**, 539 (1969).

[13] J. J. Burke and G. I. Stegeman, “Surface-polariton-like waves guided by thin, lossy metal films”, Phys. Rev. B, **33**, 5186.

[14] B. Prade, J. Y. Vinet and A. Mysyrowicz, “Guided optical waves in fibers with negative dielectric constant”, J. Lightwave Tech., **12**, 6 (1991).

[15] D. Sarid, “Long-range surface-plasma waves on very thin metal films”, Phys. Rev. Lett., **47**, 1927 (1981).

[16] H. Shi, C. Wang, C. Du, X. Luo, X. Dong and H. Gao, “Beam manipulating by metallic nano-slits with variant widths”, Opt. Express, **13**, 6815 (2005).

[17] K. Tanaka and M. Tanaka, “Simulations of nanometric optical circuits based on surface plasmon polariton gap waveguide”, Appl. Phys. Lett. **82**, 1158 (2003).

[18] G. Veronis and S. Fan, “Beam manipulating by metallic nano-slits with variant widths”, Opt. Lett., **13**, 6815 (2005).

[19] S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, J. Laluet and T. W. Ebbesen, “Channel plasmon subwavelength waveguide components including interferometers and ring resonators”, Nature **440**, 508 (2006).

[20] J. D. Jackson, *Classical Electrodynamics*, Third edition, (John Wiley & Sons, Inc. 1999).

[21] C. F. Bohren and D. R. Huffman, *Absorption and scattering of light by small particles*, (John Wiley & Sons, Inc. 1983).

[22] L. Hu, X. Chen, and G. Chen, “Surface-Plasmon Enhanced Near-Bandgap Light Absorption in Silicon Photovoltaics”, J. Comput. Theor. Naonosci. **5**, 2096 (2008).

[23] G. Laurent, N. Félidj, S. Lau Truong, J. Aubard, and G. Lévi, J. R. Krenn, A. Hohenau, A. Leitner, and F. R. Aussenegg, “Imaging Surface Plasmon of Gold Nanoparticle Arrays by Far-Field Raman Scattering”, Nano. Lett. **5**, 253 (2005).

[24] H. Xu, “Theoretical study of coated spherical metallic nanoparticles for single-molecule surface-enhanced spectroscopy”, Appl. Phys. Lett. **85**, 5980 (2004).

[25] P. Anger, P. Bharadwaj and L. Novotny, "Enhancement and quenching of single molecule fluorescence", Phys. Rev. Lett. **96**, 113002, 2006.

[26] V. M. Shalaev, *Nonlinear optics of random media*, (Springer-Verlag, 2000).

[27] U. Kreibig and M. Vollmer, *Optical properties of metal clusters*, (Springer-Verlag, 2007).

[28] K. M. Ho, C. T. Chan and C. M. Soukoulis, "Existence of a photonic gap in periodic dielectric structures", Phys. Rev. Lett. **65**, 3152 (1990).

[29] E. Yablonovitch T. J. Gmitter and K. M. Leung, "Photonic band structure: The face-centered-cubic case employing nonspherical atoms", Phys. Rev. Lett. **67**, 2295 (1991).

[30] A. Mekis, J. C. Chen, I. Kurland, S. Fan, P. R. Villeneuve and J. D. Joannopoulos, "High Transmission through Sharp Bends in Photonic Crystal Waveguides", Phys. Rev. Lett. **77**, 3787 (1996).

[31] S. Noda, M. Imada, M. Okano, S. Ogawa, M. Mochizuki and A. Chutinan, "Semiconductor three-dimensional and two-dimensional photonic crystals and devices", IEEE J. Quant. Electron, **38**, 726 (2002).

[32] H. Kosaka, T. Kawashima, A. Tomita, M. Notomi, T. Tamamura, T. Sato and S. Kawakami, "Photonic crystals for micro lightwave circuits using wavelength-dependent angular beam steering", Appl. Phys. Lett. **74**, 1370 (1999).

[33] R. Moussa, S. Foteinopoulou, Lei Zhang, G. Tuttle, K. Guven, E. Ozbay, and C. M. Soukoulis, "Negative refraction and superlens behavior in a two-dimensional photonic crystal ", Phys. Rev. B **71**, 085106 (2005).

[34] R. D. Meade, K. D. Brommer, A. M. Rappe and J. D. Joannopoulos, "Electromagnetic Bloch waves at the surface of a photonic crystal", Phys. Rev. B **44**, 10961 (1991).

[35] J. M. Elson and P. Tran, "Coupled-mode calculation with the R-matrix propagator for the dispersion of surface waves on a truncated photonic crystal", Phys. Rev. B **54**, 1711 (1996).

[36] F. Ramos-Mendieta and P. Halevi, "Surface electromagnetic waves in two-dimensional photonic crystals: Effect of the position of the surface plane", Phys. Rev. B **59**, 15112 (1999).

[37] W. M. Robertson, G.Arjavalingam, R. D. Meade, K. D. Brommer, A. M. Rappe and J. D. Joannopoulos, "Observation of surface photons on periodic dielectric arrays", Opt. Lett. **18**, 528 (1993).

[38] Y. A. Vlasov, N. Moll and S. J. McNab, "Observation of surface states in a truncated photonic crystal slab"

[39] B. Wang, W. Dai, A. Fang, L. Zhang, G. Tuttle, Th. Koschny and C. M. Soukoulis, "Surface waves in photonic crystal slabs", Phys. Rev. B **74**, 195104 (2006).

[40] H. A. Bethe, "Theory of Diffraction by Small Holes", Phys. Rev. **66**, 163(1944).

[41] T. W. Ebbesen, H. J. Lezec, H. F. Chaemi, T. Thio and P. A. Wolff, "Extraordinary optical transmission through sub-wavelength hole arrays", Nature **391**, 667(1998).

[42] , by C. Genet and T. W. Ebbesen "Light in tiny holes", Nature **445**, 39 (2007).

[43] H. F. Ghaemi, T. Thio, D. E. Grupp, T. W. Ebbesen and H. J. Lezec, "Surface plasmons enhance optical transmission through subwavelength holes", Phys. Rev. B **58**, 6779 (1998).

[44] W. L. Barnes, W. A. Murray, J. Dintinger, E. Devaux and T. W. Ebbesen, "Surface plasmon polaritons and their role in the enhanced transmission of light through periodic arrays of subwavelength holes in a metal film", Phys. Rev. Lett. **92**, 107401 (2004).

[45] T Thio, K.M. Pellerin, R. A. Linke; H. J. Lezec and T. W. Ebbesen, "Enhanced light transmission through a single subwavelength aperture", Opt. Lett. **26**,1972 (2001).

[46] H. J. Lezec, A. Degiron, E. Devaux, R. A. Linke, L. Martín-Moreno, F. J. García-Vidal and T. W. Ebbesen, "Beaming Light from a Subwavelength Aperture " Science **297**, 820 (2002).

[47] R. Moussa, B. Wang, G. Tuttle, Th. Koschny and C. M. Soukoulis, "On the beaming and enhanced transmission in photonic crystals", Phys. Rev. B **76**, 235417 (2007).

[48] E. Moreno, F. J. García-Vidal and L. Martín-Moreno, Phys. Rev. B **69**, 121402(R) (2004).

[49] P. Kramper, M. Agio, C. M. Soukoulis, A. Birner, F. Müller, R. B. Wehrspohn, U. Gösele and V. Sandoghdar, Phys. Rev. Lett. **92**, 113903 (2004).

[50] S. K. Morrison and Y. S. Kivshar, Appl. Phys. Lett. **86**, 081110 (2005).

[51] I. Bulu, H. Caglayan and E. Ozbay, Opt. Lett. **30**, 3078 (2005).

[52] W. Śmigaj, Phys. Rev. B **75**, 205430 (2007).

[53] L. Martín-Moreno, F. J. García-Vidal, H. J. Lezec, A. Degiron and T. W. Ebbesen, Phys. Rev. Lett. **90**, 167401 (2003).

[54] D. Z. Lin, C. K. Chang, Y. C. Chen, D. L. Yang, M. W. Lin, J. T. Yeh, J. M. Liu, C. H. Kuan, C. S. Yeh and C. K. Lee, Opt. Express **14**, 3503 (2006).

[55] F. J. García-Vidal, L. Martín-Moreno, H. J. Lezec and T. W. Ebbesen, Appl. Phys. Lett. **83**, 4500 (2003).

[56] L.-B. Yu, D.-Z. Lin, Y.-C. Chen, Y.-C. Chang, K.-T. Huang, J.-W. Liaw, J.-T. Yeh, J.-M. Liu, C.-S. Yeh and C.-K. Lee, Phys. Rev. B **71**, 041405(R) (2005).

[57] G. Gay, O. Alloschery, B Viaris De Lesegno, C. O'Dwyer j. Weiner and H. J. Lezec, Nature Phys. **2**, 262(2006).

[58] P. Lalanne and J. P. Hugonin, Nature Phys. **2**, 551(2006).

[59] L. Chen, J. Robinson and M. Lipson, Opt. Express 14, 12629 (2006).

[60] L. Aigouy, P. Lalanne, J. P. Hugonin, G. Julíe, V. Mathet and M. Mortier, Phys. Rev. Lett. **98**, 153902(2007).

[61] X. Y. Yang, H. T. Liu and P. Lalanne, Phys. Rev. Lett. **102**, 153903(2009).

[62] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic, New York, 1985).

[63] F. J. García-Vidal and L. Martín-Moreno, Phys. Rev. B **66**, 155412(2002)

[64] P. Lalanne, J. P. Hugonin and J. C. Rodier, Phys. Rev. Lett. **95**, 263902(2005).

[65] F. J. García-Vidal, H. J. Lezec, T. W. Ebbesen and L. Martín-Moreno, Phys. Rev. Lett **90**, 213901(2003).

[66] J. A. Porto, F. J. García-Vidal and J. B. Pendry, Phys. Rev. Lett. **83**, 2845(1999).

[67] Q. Cao and P. Lalanne, Phys. Rev. Lett. **88**, 057403(2002).

[68] K. G. Lee and Q. Han Park, Phys. Rev. Lett. **95**, 103902(2005).

[69] O. T. A. Janssen, H. P. Urbach and G. W. 't Hooft, Phys. Rev. Lett. **99**, 043902(2007).

[70] Lord Rayleigh, "On the Dynamical Theory of Gratings", Proc. Roy. Soc. (London) **A79**, 399 (1907); "Note on the remarkable case of diffraction spectra described by Prof. Wood", Phil. Mag. **14**, 60 (1907).

[71] J. E. Stewart and W. S. Gallaway, "Diffraction anomalies in grating spectrophotoneters", Appl. Opt. **1**, 421 (1962).

[72] K. J. Klein Koerkamp, S. Enoch, F. B. Segerink, N. F. van Hulst and L. Kuipers "Strong influence of hole shape on extraordinary transmission through periodic arrays of subwavelength holes", Phys. Rev. Lett. **92**, 183901 (2004).

[73] Z. Ruan and M. Qiu, "Enhanced Transmission through Periodic Arrays of Subwavelength Holes: The Role of Localized Waveguide Resonances", Phys. Rev. Lett. **96**, 233901 (2006).

[74] P. Markos and C. M. Soukoulis, *Wave Propagation: From Electrons to Photonic Crystals and Left-handed Materials* (Princeton University Press, New Jersey, 2008).

[75] S. Kim, H. Kim, Y. Lim and B. Lee, Appl. Phys. Lett. **90**, 051113(2007).

[76] H. Caglayan, I. Bulu and E. Ozbay, J. Appl. Lett. **104**, 073108 (2008).

[77] H. Caglayan, I. Bulu and E. Ozbay, Appl. Phys. Lett. **92**, 092114 (2008).

[78] E. Ozbay, Science, 311, 5758(2006).

[79] S. Kim,, Y. Lim, H. Kim J. Park and B. Lee, Appl. Phys. Lett. **92**, 013103(2008).

[80] O. T. A. Janssen, H. P. Urbach and G. W. 't Hooft, Phys. Rev. Lett **99**, 043902(2007).

[81] N. Yu, J. Fan, Q. Wang, C. Pflüg, L. Diehl, T. Edamura, M. Yamanishi, H. Kan and F. Capasso, Nature Photon. **2**, 564(2008).

[82] A. Taflove and S. C. Hangess, Cimputational electrodynamics: the finite-difference time-domain method, Second edition, Artech House Publishers (2000).

[83] Z. S. Sacks, D. M. Kingsland, R. Lee and J. F. Lee, "A Perfect matched anisotropic absorber for use as an absorbing boundary condition", IEEE Trans. Antennas and Propagation, **43**, 1460 (1995).

[84] M .Koshiba, Y. Tsuji and S. Sasaki, "High-Performance Absorbing Boundary Conditions for Photonic Crystal Waveguide Simulations", IEEE Microwave Wireless Comput. Lett. **11**, 152 (2001).

[85] E. P. Kosmidou, T. I. Kosmanis and T. D. Tsiboukis, "A comparative FDTD Study of Various PML Configurations for the Termination of Nonlinear Photonic Bandgap Waveguide Structures", IEEE Trans. Magn. **39**, 1191 (2003),

AUTHOR PUBLICATIONS

1. B. Wang, W. Dai, A. Fang, L. Zhang, G. Tuttle, T. Koschny and C. M. Soukoulis, "Surface waves in photonic crystal slabs", *Phys. Rev. B* 74, 195104 (2006).

2. W. Dai, B. Wang, T. Koschny and C. M. Soukoulis, "Experimental verification of quantized conductance for microwave frequencies in photonic crystal waveguides", *Phys. Rev. B* 78, 073109 (2008).

3. W. Dai and C. M. Soukoulis, "Converging and wave guiding of Gaussian beam by two-layer dielectric rods", *Appl. Phys. Lett.* 93, 201101 (2008).

4. W. Dai and C. M. Soukoulis, "Theoretical analysis of the surface wave along the metal-dielectric interface", accepted by *Phys. Rev. B*.

5. W. Dai and C. M. Soukoulis, "Control of beaming angles via a subwavelength metallic slit surrounded by grooves", submitted to *Phys. Rev. B*.