

Optical and Semiconductor Devices Group
Department of Electrical and Electronic Engineering
Imperial College London

Metrology of High Aspect Ratio Nanoelectromechanical Systems (NEMS)

By: Fu Yee Kwan
Supervised by: Prof. Richard Symms and Dr Oleksiy Sydoruk

A thesis submitted in fulfillment for the degree of
Doctor of Philosophy

November 2023

Declaration of Originality

I certify that all material in this thesis entitled, 'Metrology of High Aspect Ratio Nanoelectromechanical Systems (NEMS)', is my original own work under the supervision of Professor Richard Syms and Dr Oleksiy Sydoruk. Other material that is not my own has been appropriately referenced and acknowledged.

Fu Yee Kwan

Declaration of Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

This thesis investigates the possibility of substituting optical imaging of nano-electro-mechanical systems (NEMS) for the lengthy and costly process of scanning electron microscope inspection. Although image details are both modified and degraded by diffraction, nanoscale features can be seen in experimental images from conventional light microscopes, and excellent contrast is achieved in dark field (DF). The motivation is therefore to analyse scattering from sub-wavelength structures with visible electromagnetic (EM) waves and to establish a simple method of simulating their DF images, an otherwise extremely lengthy process.

A comprehensive review of different theories of rigorous two-dimensional (2D) diffraction theory is then presented. Geometrical theory of diffraction (GTD) is chosen as the best approach and validated by numerical simulation of the imaging of a set of periodic 2D canonical structures using the finite element method, which is compared with modal diffraction theory. It is demonstrated that only edge waves contribute to dark field images, and that GTD wedge diffraction coefficients provide good approximations to their amplitudes. A 2D model is then developed to account for direct scattering and re-scattering of in-plane radiation. The model is valid for both TE and TM polarizations.

Scattering from simulated optical imaging of 3D canonical structures is then investigated. The numerical results are carefully considered to establish the minimum number of waves needed to simulate diffuse and polychromatic illumination. A novel simulation tool is then developed to reduce the time taken to simulate dark field imaging of general NEMS. This model uses edge detection to identify the scattering edges, a convolution method to simulate in-plane scattering, and wedge diffraction coefficients from GTD to define edge wave amplitudes. The results are compared with 3D FEM simulations, and excellent agreement is obtained. Deficiencies in the model are identified and suggestions for future developments are made.

Acknowledgments

I would like to thank my supervisor Professor Richard Syms and my co-supervisor Dr. Oleksiy Sydoruk, for their guidance, support, and encouragement during the past five years. This work could not have been completed without their unlimited support and valuable suggestions. Their constructive criticisms have been critical to my writing skills and technical understanding.

I also wish to thank Dr. Simone Zonetti and Dr. Attique Dawood, former members of the Optical and Semiconductor Devices group at Imperial College London, for providing technical support about the simulation tool COMSOL Multiphysics. Also, I want to express my gratitude to everyone in room 707 for the fruitful discussions on various topics and suggestions to my work.

I am grateful to my parents for their mental and financial support to allow me to pursue a PhD degree, and my brother for encouraging me to explore my interests in engineering since I was young. Last but not least, I would like to thank my partner, Arisa for her love, support and understanding so that I can focus on my research in the past years.

Contents

Declaration of Originality	ii
Declaration of Copyright	iii
Abstract	iv
Acknowledgements	v
Chapter 1 Introduction	1-1
1.1 MEMS and NEMS devices	1-1
1.2 MEMS and NEMS fabrication	1-4
1.3 MEMS and NEMS metrology	1-11
1.4 Thesis overview	1-14
References	1-17
Chapter 2 Theoretical Background	2-1
2.1 Maxwell's equations	2-1
2.2 Electromagnetic waves	2-3
2.3 Fourier optics	2-6
2.4 Optical microscopy	2-7
2.5 Image formation	2-9
2.6 Resolution	2-11
2.7 HAR MEMS and NEMS objects	2-13
2.8 Summary	2-15
References	2-16
Chapter 3 Diffraction Theory of Isolated Obstacles	3-1
3.1 Sommerfeld's half-plane diffraction problem	3-1
3.2 Geometrical theory of diffraction (GTD)	3-6
3.2.1 Fresnel integral	3-10
3.2.2 Uniform theory of diffraction (UTD)	3-14
3.3 Physical theory of diffraction (PTD)	3-16
3.4 Incremental theory of diffraction (ITD)	3-18
3.5 Multiple diffraction	3-20

3.6 Summary	3-22
References	3-23
 Chapter 4 Diffraction from an Isolated Structure	 4-1
4.1 Analytical model	4-1
4.2 Finite element modelling	4-3
4.3 Diffraction from a wedge	4-6
4.3.1 The nature of cylindrical wave	4-7
4.3.2 Diffraction coefficients	4-8
4.4 Diffraction from a mesa and a half-plane	4-12
4.5 TM polarization	4-15
4.6 Summary	4-19
References	4-20
 Chapter 5 Diffraction from Periodic Structures	 5-1
5.1 Modal diffraction model	5-1
5.1.1 Single mesa	5-2
5.1.2 Double mesa	5-5
5.2 Periodic finite element model	5-6
5.2.1 Back-scattered field (E_S)	5-7
5.2.2 Amplitude of diffraction orders (r_L)	5-8
5.2.3 Filtering of diffraction orders	5-9
5.2.4 Image formation	5-10
5.3 Numerical results	5-11
5.4 Modifications	5-16
5.4.1 Modified model	5-16
5.4.2 Single mesa diffraction	5-17
5.4.3 Double mesa diffraction	5-21
5.5 Summary	5-23
References	5-25
 Chapter 6 Simulation of 3D Images	 6-1
6.1 Scattering from 3D structures	6-1

6.2 Comparison with 2D model	6-4
6.3 Simulation of 3D imaging	6-4
6.3.1 Rectangular structures	6-5
6.3.2 Curved structures	6-7
6.4 Image library	6-9
6.4.1 Single mesas	6-9
6.4.2 Double mesas	6-10
6.4.3 Crossed mesas	6-11
6.4.4 Double crossed mesas	6-12
6.4.5 Star patterns	6-12
6.4.6 Pillars and cylinders	6-13
6.4.7 Finite structures	6-14
6.4.8 Indeterminate cases	6-15
6.5 Additional out-of-plane illumination angles and wavelengths	6-18
6.5.1 Multiple out-of-plane illumination angles	6-18
6.5.2 Polychromatic illumination	6-20
6.6 TM polarization	6-21
6.7 Summary	6-23
References	6-24
 Chapter 7 Novel Analytical Simulation Tool	 7-1
7.1 Introduction	7-1
7.2 2D model – TE polarization	7-3
7.3 2D model – TM polarization	7-11
7.4 Huygens' principle and the ITD	7-14
7.5 3D model	7-16
7.6 Numerical results	7-19
7.6.1 Canonical objects	7-20
7.6.2 NEMS devices and cellular materials	7-21
7.6.3 Further limitations	7-22
7.6.4 TM incidence	7-23
7.7 Summary	7-23
References	7-36

Chapter 8 Summary and Future Work	8-1
8.1 Conclusions	8-1
8.2 Future work	8-3
8.2.1 Model improvements	8-3
8.2.2 Comparison with experiment	8-4

1. Introduction

The aim of this work is to develop methods of simulating optical images of micro- and nano-electro-mechanical systems (MEMS and NEMS), as a step towards validating the use of low-cost optical metrology for NEMS. Overviews of devices, fabrication processes and current metrology will therefore be given, followed by an outline of the thesis.

1.1 MEMS and NEMS devices

The idea of miniaturisation of electronics can be dated back to the mid 20th century, when the first integrated circuit (IC) chip was created. The process of combining metal-oxide-semiconductor (MOS) transistors onto a single chip, known as large-scale integration (LSI), began in the following years. Initially, the numbers of transistors were small (thousands); today they are measured in billions. The MOS IC chip is the mainstay of the semiconductor industry especially for processor and memory applications.

As micromachining and packaging techniques matured in the 1970s, applications beyond microelectronics were explored. A key driver was an early observation of the excellent mechanical properties of silicon [1] and inspiring demonstration of movable microstructures [2]. Microsensors were first commercialised in the automotive industry [3], and automotive applications have gradually spread from safety systems to ride control [4]. With the emergence of the surface micromachining in the 1980s [5], novel sensors and actuators became increasingly practical [6] and applications became widespread [7].

Microscale integrated devices or systems, which consist of both electrical and mechanical components, form the basis of microelectromechanical systems (MEMS). Component sizes usually range from a few to over hundred micrometres. General mechanical features include supports with different degrees of freedom, such as a built-in or encastré supports, together with joints and loads. Different support arrangements allow linear and rotary motion, and exploitation of many common elastic effects, including beam bending (for suspension systems) and buckling (for bistable devices). Often, the dynamics of vibrating MEMS are modelled as classic mass-spring-damper systems, but advantage is also taken of the full mode spectrum offered by vibrating beams, plates, and shells.

In contrast to LSI, the level of integration is extremely low, and MEMS typically consist of single devices. However, miniaturisation has brought numerous advantages over conventional devices: lower power consumption, increased sensitivity and accuracy, improved repeatability and reliability, lower cost, and ease of integration into systems. Applications are not restricted to electrical and mechanical devices, but include power MEMS [8], RF MEMS [9], optical MEMS [10], ion optical MEMS [11], microfluidics [12], medical MEMS [13] and bioMEMS [14]. In each case, specialised materials, device formats and packages are needed. Devices have impacted the aerospace, automotive, communications, defence, IT, metrology, medical and pharmaceutical industries.

As their name suggests, MEMS usually consist of a combination of electrical and mechanical parts and use transducers to transform energy between electrical and mechanical form. For example, a microactuator can be driven by electrical signals to generate motion. Actuators are commonly used to excite motion in resonant sensors, deflect mirrors, control fluidic valves or switch electrical contacts. There are four main types: electrostatic, electrothermal, piezoelectric and electromagnetic [15, 16]. All obey different size scaling laws, with electromagnetic being the least effective. Electrostatic and electrothermal actuators are the simplest and are widely used in MEMS.

Electrostatic actuators use the electrostatic force between the plates of a charged capacitor to create motion in the desired direction, either out-of-plane or in-plane. Figure 1.1 shows the two main geometries of electrostatic actuators: a) parallel plate and b) comb-drive. The parallel plate arrangement is designed for out-of-plane movement and the comb-drive for in-plane movement. The black features are the MEMS parts, and the grey surround is the substrate. The dotted lines show the idle position of the moving part. When a voltage is applied, the suspended parts will move in the same directions.

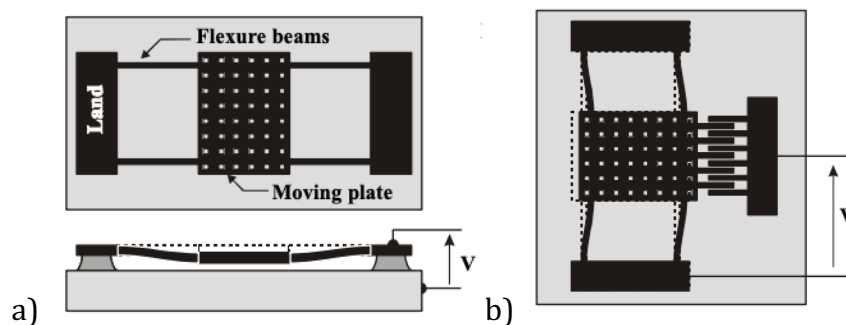


Figure 1.1: Electrostatic actuators with a) parallel plate and b) comb-drive electrodes.

The electrothermal actuator is another popular type, due to its simplicity. The simplest type – buckling mode actuators – exploit the temperature rise obtained by passing an electrical current through a deposited thin-film heater or through the mechanical layer to generate thermal expansion, which is then converted into motion by constraining the expansion. Bimorph actuators rely on differential thermal expansion between adjacent parts of a structure. They can be realised by layering together two materials with different coefficients of thermal expansion, or by exploiting layout to cause differential heating to occur in different regions of a structure formed in a common material. Figure 1.2 shows the arrangements of a) material bimorph and b) shape bimorph actuators. The principles of both are the same: a beam bending effect due to the difference in expansion between the two materials or arms.

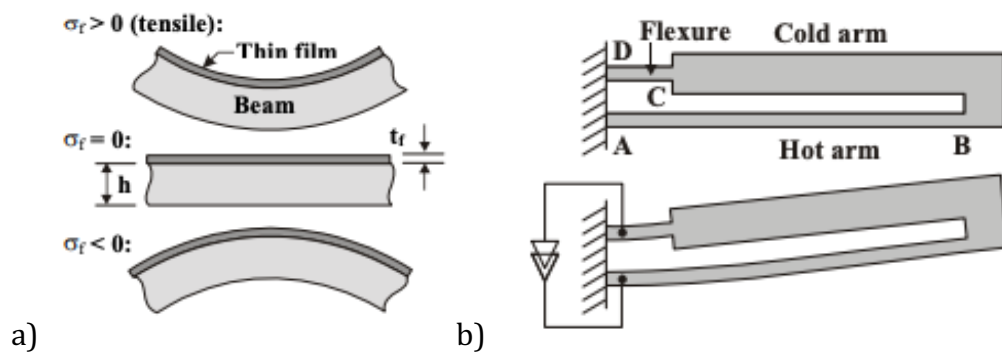


Figure 1.2: Electrothermal actuators: a) material and b) shape bimorph.

A MEMS device can also be used as a sensor to convert motion to electrical current. An electrostatic sensor measures the in-plane or out-of-plane movement by the change of capacitance between the plates of the capacitor, with the same geometry as Figure 1.1. For a piezoelectric sensor, it measures the dielectric polarisation and the surface charge generated from external pressure. Inertial sensors such as accelerometers or gyroscopes are used in many MEMS applications, for example, automotive crash-bag and ride control systems. The external acceleration applied to an accelerometer causes an internal mass displacement which can be converted to electrical current by a transducer. Similarly, gyroscopes measure rotational acceleration by exploiting Coriolis acceleration.

Around the 2000s, researchers started investigating the submicron regime. The resulting devices are known as nanoelectromechanical systems (NEMS). Apart from the size reduction, down-scaling has enabled higher mechanical responsivity at higher fundamental frequencies resulting in larger force sensitivity [17]. The importance of

quantum effects (such as zero point motion) and of high surface-to-volume ratio [18] have made NEMS more useful for surface-based sensing [19]. Reduced mass and higher electrical strength are further benefits [20, 21]. However, NEMS do not replicate MEMS at nanoscale. Particularly, their electrical, and mechanical behaviours and material properties may be quantitatively different [22]. Major challenges for commercialisation of NEMS also lie in device fabrication and metrology, as described below [23].

1.2 MEMS and NEMS fabrication

In principle, integrated circuits, MEMS, and NEMS can all be fabricated using three primary processes: pattern transfer, material deposition and material removal [24]. MEMS devices are commonly fabricated in foundries [25-27] that are similar in concept to IC fabs, offering a range of standard processes that can be combined to form a variety of different device types by following simple design rules [28]. However, the processes are often specialised [29-31], and the materials systems for high-value Si MEMS (which often require co-integration with electronics) remain stubbornly different to those needed for other MEMS (for example, microfluidics, which require disposable plastics).

Very high production rates, high yields and low costs are obtained by replication of a set of master mask patterns on substrates or wafers. Device operations and process flows are first simulated [32-35], generally using commercial finite element modelling (FEM) software such as COMSOL® [36] or ANSYS® [37]. The masks are then designed using standard layout software and fabricated by direct writing using an electron-beam or laser system. For ICs, many masks are needed, and a complete set for large wafers containing additional features needed for resolution enhancement can therefore be extremely expensive. Integrated circuit layouts typically follow ‘Manhattan’ geometries, while MEMS and NEMS layouts can vary widely. For example, Figure 1.3a) shows an mask pattern containing electrostatic MEMS comb-drives and additional NEMS elements

For ICs, pattern transfer is carried by photolithography, using a stepping exposure system based on projection images of mask patterns [38] generated at very short wavelength [39] and high numerical aperture [40]. Photolithography is a parallel process with high throughput. Despite considerable development, feature size is currently restricted to 65 nm or larger [41] although this may reduce [42]. Steppers are extremely

expensive, and their use is generally unjustified for large area, low-cost devices such as MEMS or NEMS. Lower cost contact aligners operating with mid-UV light are therefore often used. For example, Figure 1.3b shows a contact mask aligner in use at Imperial College London, which also allows double sided patterning. However, conventional aligners are (at least, nominally) incapable of transferring nanoscale patterns. Instead, NEMS require advanced lithography methods. Because of this limitation, there is no real equivalent of a NEMS foundry, and few devices have been commercialised; the most common are sharp tips for scanning probe microscopy [43, 44].

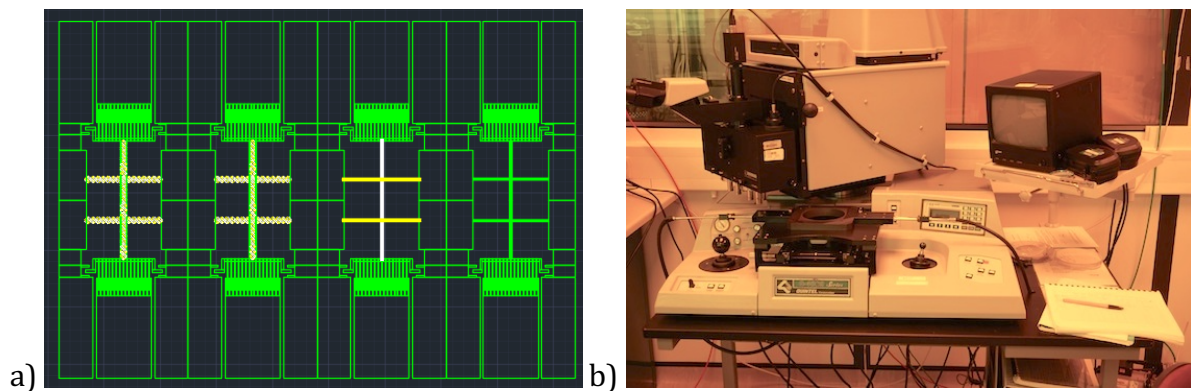


Figure 1.3: a) Example mask pattern for electrostatic MEMS, and b) mid-IV Quintel Q-4000 contact mask aligner (Courtesy Prof. R. Syms and Dr D. Liu).

Photolithography is based on processes developed for integrated circuits [45-47]. Contamination is first removed and chemicals such as hexamethyl disilazane (HMDS) are used to promote resist adhesion. The wafer is then covered with a thin, uniform photoresist layer by spin coating (possibly using a specialist resist designed for specialised MEMS applications [48, 49]). Soft baking at 100°C is used to drive off residual solvent and improve adhesion. The photoresist is then exposed and developed. Exposure to UV light transfers a master pattern from a mask onto the resist as a latent image. Post-exposure baking is used to reduce standing wave phenomena caused by interference of the incident light. The latent image is converted to a topographic pattern using a developer such as tetramethyl ammonium hydroxide (TMAH). Hard baking at a higher temperature is then performed to improve chemical and physical robustness of the patterned resist layer. The pattern is then solid enough for etching or ion implantation.

For smaller structures, next generation lithography (NGL) is needed for non-optical lithography, such as extreme UV (EUV) lithography [39] and nanoimprint lithography [50-52]. Among the non-optical lithography methods, electron beam lithography (EBL) is the best option because it allows rapid prototyping with extremely high resolution [53]. To write a pattern, a focused electron beam is simply scanned over the resist surface using an electrostatic deflector. The resist is modified by the electrons (for positive resist, by breaking down the polymer chains). The resolution is approximately 25nm, depending on the spot size, lens aberrations and resist. However, research has shown that e-beam lithography is capable of patterning feature widths less than 10nm over 300mm diameter substrates. The fundamental limit to resolution is electron scattering [54, 55]

EBL requires expensive capital equipment and involves extremely long serial writing and is consequently unsuitable for volume production of low-cost devices. Other nanoscale patterning methods such as scanning probe lithography [56-58], focused ion beam lithography [59, 60] and dip pen nanolithography [61, 62] are also serial and unsuitable for high-volume production due to their low throughput. Parallel patterning methods, such as nanosphere lithography [63, 64] or self-organising resists [65-67], can only form a highly restricted range of features, often in periodic or random arrangements; these are suitable for specialised surfaces but not for defined device layouts.

Once the resist is patterned, the next step is to transfer the pattern into the substrate, typically by chemical removal of material. The most common techniques are wet etching and deep reactive ion etching (DRIE). Wet etching uses strong acids, and its advantages are high throughput and high selectivity. However, etching is generally isotropic [68]. For silicon, anisotropic etching terminating on crystal planes can be achieved with specific wet etchants [69]. In this case, the process is commonly known as bulk micromachining [70]. However, the range of structures that can be fabricated in this way is very restricted. In surface micromachining, isotropic etching of sacrificial materials such as silica is often used to undercut suspended structures [31]. Isotropic etching of SiO_2 and Si can also be achieved using vapor phase HF and XeF_2 respectively [71, 72]. Gas phase processes of this type avoid collapse of suspended structures due to surface tension.

Deep reactive ion etching is an anisotropic process [73, 74], which can form high aspect ratio structures with excellent sidewall verticality and quality using a combination of plasma etching and sidewall passivation, often in a cyclical process [75, 76]. In the first step, a reactive gas is excited into a plasma using RF energy, often introduced inductively. Mobile ions are then accelerated towards the surface of the wafer, where they form a volatile by-product that may easily be pumped away. The exposed sidewall is then passivated with a polymer. A typical DRIE process (for example, the Bosch process) uses SF_6 to provide the reactive species for etching Si and C_4F_8 to provide the passivating layer [73]. Figure 1.4a shows a DRIE system at Imperial College London, and Figure 1.4b shows a scanning electron microscope view of high aspect ratio (HAR) silicon nanostructures fabricated using the Bosch process by Dr Dixi Liu.

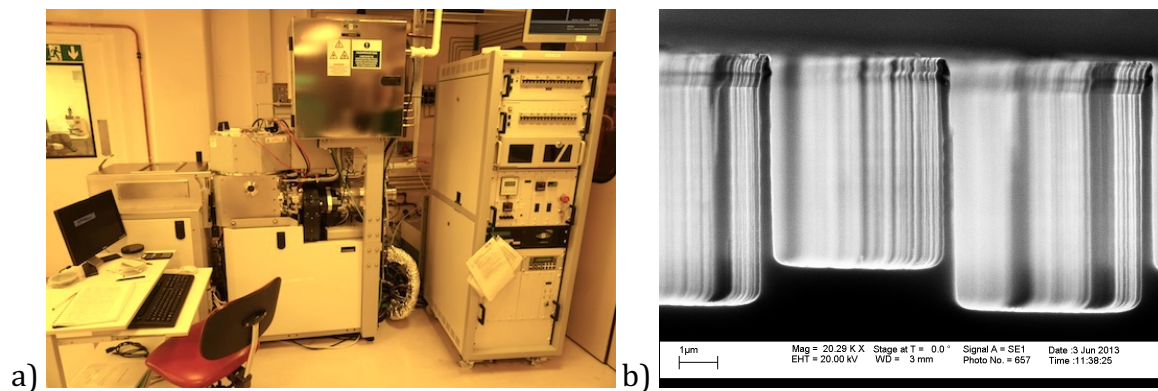


Figure 1.4: a) STS Single chamber Multiplex DRIE system; b) deep etched silicon nanostructures (Courtesy Prof. R. Syms and Dr D. Liu).

Deposition is used to form layers of silicon and non-silicon materials that can provide different parts of the device structure. The most common deposition technologies for semiconductors and dielectrics are chemical vapour deposition (CVD), molecular beam epitaxy (MBE) and liquid phase epitaxy (LPE). Amorphous materials such as SiO_2 and Si_3N_4 are deposited using low-pressure (LPCVD) and plasma enhanced (PECVD) chemical vapour deposition [77, 78]. In LPCVD, the substrate is placed in a furnace tube supplied with precursor gases, which react to produce solid material by condensation on the wafer. LPCVD produces layers with excellent uniformity of thickness and material characteristics. However, the deposition temperature is over 600°C and the resulting films are too highly stressed for MEMS or NEMS. PECVD operates at lower temperature due to the extra energy supplied by a plasma. Consequently, intrinsic stresses are lower, but the deposited films are less uniform, and often relatively porous [79].

Epitaxy is a similar process to CVD but allows the deposition of crystalline semiconductor material on a crystal [80]. Vapour phase epitaxy (VPE) is the most common method, and single crystal and polycrystalline materials can both be deposited [81, 82]. Unfortunately, the precursor gas for silicon (silane, SiH_4) is highly dangerous. An advantage of epitaxy is the high growth rate, and the thickness of the film can be over $100\mu\text{m}$. Single crystal material cannot be grown on amorphous layers such as SiO_2 . However, silicon-on-insulator (SOI) substrates can be formed by implantation of oxygen ions to form a thin buried SiO_2 layer capped with a thin Si layer [83]. An alternative, bonded silicon-on-insulator (BSOI), simply involves bonding an oxidised silicon wafer to a bare wafer, which is then polished back to any desired (and larger) thickness [83, 84]. Both methods provide high quality Si device layers on top of a sacrificial SiO_2 layer, and hence form a simple route to fabricating suspended movable parts by undercut etching.

Several different processes are used for metal deposition: CVD, thermal evaporation, sputter coating, and electroplating. CVD has already been described. In thermal evaporation, the substrate is placed in a vacuum chamber with a small quantity of source material. When this material is heated to its evaporation point, source atoms condense on the substrate surface to form thin films [86]. Heating methods include nonreactive metal filaments and electron beams. Because of the temperatures involved, evaporated films have high intrinsic stress. RF sputtering operates at a lower temperature. A substrate and a planar target formed from pure source material are mounted facing each other inside a vacuum chamber. A non-reactive gas plasma (for example, argon) is generated by a RF power source and ions are accelerated towards the target. Target atoms are physically sputtered from the target and condense on the substrate surface [87]. RF sputtering generates conformal films that are thicker than those achievable by evaporation [88]. However, electroplating is required to achieve very thick films.

In electroplating, a sputtered seed layer is first deposited. The substrate is placed in a liquid electrolyte containing a suitable source of material to be deposited (for example, CuSO_4 for copper, combined with H_2SO_4 to increase conductivity and Cl_2 to control plating). When a voltage is applied between the substrate and a counter electrode, a redox reaction takes place to deposit a layer of material. Electroplating deposits low-stress material [89] and may be carried out on a material that has already been patterned with

a resist, so that the deposited material is formed inside a mould [90]. Excess seed layer can then be removed. Because it is a low-temperature post-process, electroplating provides a simple route to co-integration with electronics [91].

Many processes have been developed to form MEMS devices, using combinations of lithography, etching and material deposition. Often these are used repetitively, allowing successive layers of materials to have different properties (for example, high or low etching rates, high mechanical strength, or high electrical conductivity). Processes may be additive (or bottom-up) or subtractive (top-down). Similar approaches may be taken to NEMS fabrication, and a variety of processes have been developed to etch high-aspect ratio nanostructures, for example using electrochemical or photoelectrochemical etching [92-94]. However, the key difficulty in NEMS fabrication, which has already been mentioned, is the need for high-resolution lithography.

One possible solution is sidewall transfer lithography (STL), originally developed by the IC industry to form FINFETs [95] and nanowires [96, 97] and adapted at Imperial College to form HAR MEMS with nanoscale features such as flexible Si suspensions [98]. The process can be carried out using low-cost optical lithography and deep reactive ion etching. Figure 1.5a shows the principle. The wafer is patterned using Mask 1, the perimeters of whose features are used to define nanoscale parts. The wafer is then patterned using Mask 2, whose features are used to define microscale parts. The two patterns are then transferred into the substrate together by DRIE, to yield a 3D structure containing features with combined size scales as shown in Figure 1.5b.

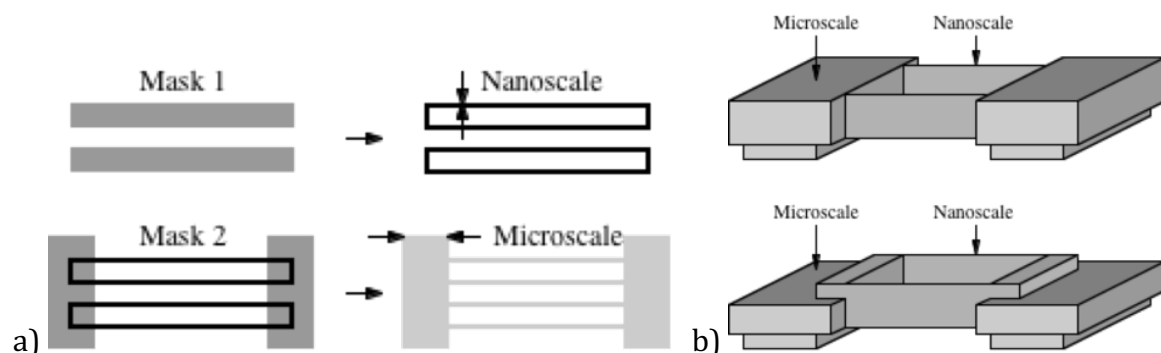


Figure 1.5: a) Principle and b) result of STL NEMS processing (courtesy Prof. R. Syms).

Figure 1.6a shows a section view of a STL fabrication process for BSOI NEMS. The grey and white layers are Si and SiO₂ respectively; the grey layers are photoresist and the two darkest layers are chromium and gold (Cr/Au) and aluminium (Al). First, the device layer is patterned with resist and a mask that defines the layout of nanoscale features, using conventional lithography (1). DRIE is used to transfer the pattern into the silicon as shallow mesas (2) and the resist is stripped (3). Next, RF sputtering is used to coat the mesas with a conformal layer of metal (4). Directional etching is then used to remove the horizontal metal surfaces, leaving the metal on vertical surfaces (5). The width of the vertical features is defined by the original coating thickness, but not by the optical lithography. The spacing between the vertical features is the width of the mesa. The second lithography step is taken place to pattern the structure once more with a mask that defines the microscale features (6). DRIE is used again for pattern transfer into the silicon device layer (7), and resist is removed (8). HF vapour undercut is then used to remove the oxide interlayer (9). The nanoscale part A is suspended and the microscale part B acts as an anchor. Finally, the entire structure is metallised with aluminium (10) to improve electrical contact. Figure 1.6b shows an SEM view of a completed device fabricated by Dr Liu, showing MEMS and NEMS parts suspended above a substrate.

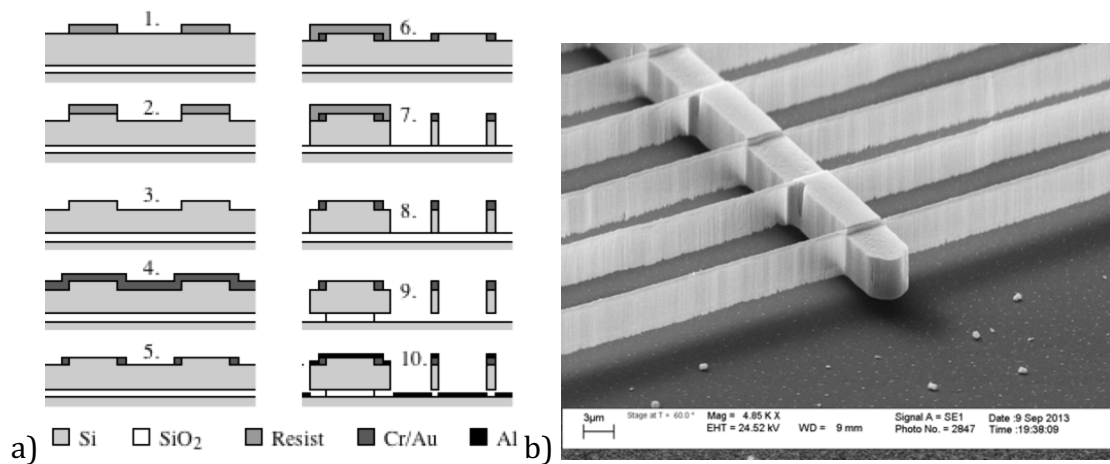


Figure 1.6: a) Process flow for BSOI NEMS fabrication; b) SEM view of completed device (Courtesy Prof. R. Symms and Dr D. Liu).

The process has several limitations. The nanoscale features must all follow polygonal perimeters, and (without further processing) these must be non-intersecting. Furthermore, the spacing between nanoscale features is normally set by the resolution of the lithography and is therefore typically microscale. However, if the STL patterning step

is performed multiple times, intersecting nanoscale features can be formed [99], and complex actuators and nanostructured cellular materials [100] have already been demonstrated. STL does not require a nanostructured master or high-resolution lithography as the width of the nanofeature depends only on the thickness of the coating. Hence, it is a low-cost process that may allow foundry scale NEMS production.

Metrology of micron-scale structures can easily be performed by including conventional microscopy. Considerable attention has been paid to optical metrology of simple structures such as resist lines for integrated circuit fabrication, and methods were developed for simulation of line images [101-106]. Generally, these combined accurately calculated diffracted fields obtained using modal theory [107-109] or rigorous coupled wave theory [110, 111] with Fourier-domain models of microscopy [112-115]. Consequently, simulation was mainly limited to periodic structures in 2D. For example, Figure 1.7a shows the feature layout of a periodic pattern of resist stripes in an early simulation [102], and Figure 1.7b compares the simulated and experimentally measured line-scan showing additional edge features introduced by diffraction. This research is directly relevant to the current work and will be considered in more detail later.

Figure 1.7: a) Feature layout for simulation of resist images; b) simulated linescan (after Nyysönen [102]).

profile of a complex multilayer surface mask for a device with a microscale comb-drive and a nanoscale cellular suspension; here the outline and relative thickness of each layer can clearly be seen, and the result of subsequent etching can be easily inferred.

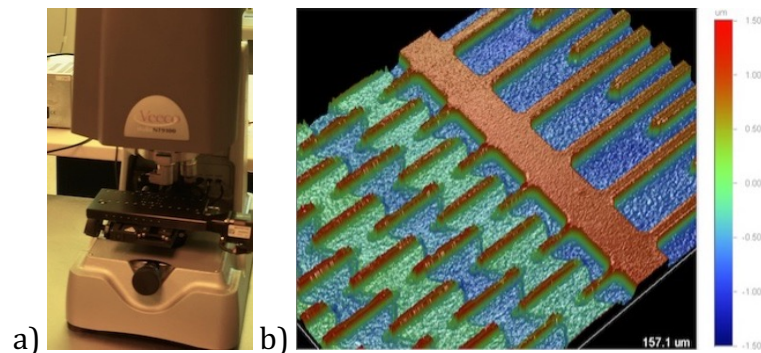


Figure 1.8: a) Veeco NT9100 Surface Profilometer and b) surface profile of MEMS/NEMS device (courtesy Prof. R. Syms and Dr D. Liu).

SEM inspection remains the gold standard. For example, Figure 1.9a shows the LEO 1450 VP SEM currently in use at Imperial College London, and Figure 1.9b shows an SEM image of optical MEMS device combining deep-etched electrostatic comb drives with a surface relief diffraction grating. The image combines fine detail, high contrast and a 3D perspective. However, operation of SEM requires a vacuum environment, and imaging is a slow serial process that limits throughput. Electron bombardment can also heat, charge, and deflect sensitive suspended parts, the essence of MEMS and NEMS.

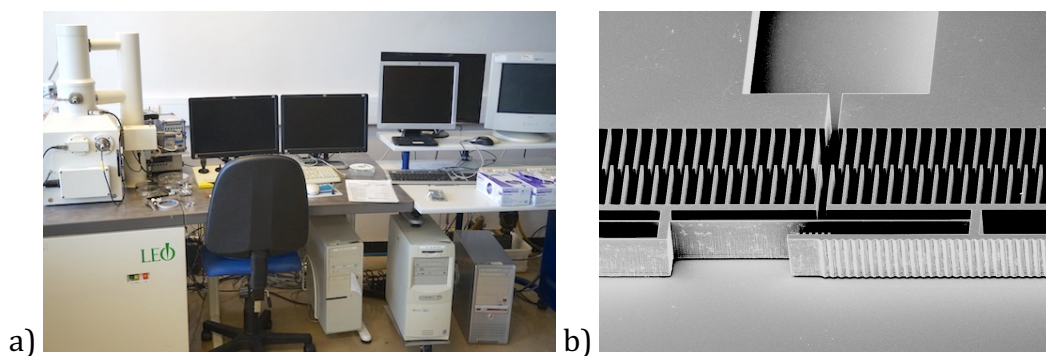


Figure 1.9: a) LEO 1450 VP SEM and b) SEM image of optical MEMS device (courtesy Prof. R. Syms).

Metrology is considerably more difficult for NEMS and represents another factor preventing commercialisation. Sub-micron linewidths can be measured using interferometric microscopes [121, 122], and, despite nominally insufficient resolution, strongly submicron features can be visualised in an optical microscope image [123, 124].

For example, Figure 1.10a shows bright field (BF) images of suspended NEMS parts attached to a mesa. The width, depth and spacing of all nanoscale parts is $0.1\ \mu\text{m}$, $4\ \mu\text{m}$, and $10\ \mu\text{m}$, respectively. The nanoscale parts are seen as dark lines on the left and the mesa as the grey area with a dark edge on the right. The bright background is the Si substrate. Figure 1.10b shows the corresponding dark field (DF) image. The nanoscale parts now appear as bright lines, and the dim vertical line represents the edge of the mesa, which is now dark. Although the resolution is lower than in SEM images and the presentation is 2D, it is clearly possible to distinguish NEMS and MEMS parts.

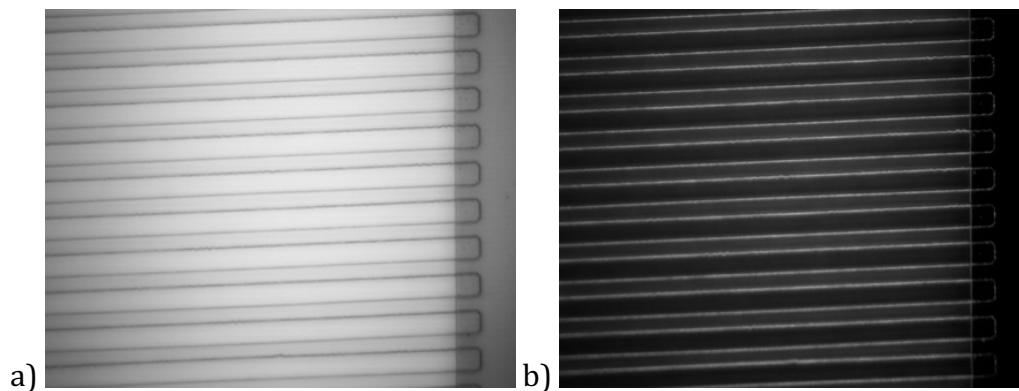


Figure 1.10: Optical images of a suspended NEMS attached to a mesa in a) bright and b) dark field illumination (courtesy Prof. R. Syms and Dr D. Liu).

Figure 1.11 shows a more complex MEMS/NEMS combined device fabricated in bulk Si by multilayer STL. The device is an electrothermal microgripper, an important tool in biomedical industry for applications such as cell manipulation. Figure 1.11a shows a BF image of the grips, with cold arms on the outside and hot arms on the inside. The difference in brightness between the micro and nano parts is again significant; the grey regions are microscale parts, and the dark lines are the nanoscale parts of the gripper arms. The bright background is the Si substrate. Figure 1.11b shows the corresponding DF image. The dim lines now show the outline of the microgripper and the cold arms. The bright lines are the nanoscale hot arms. For the flexure between the cold arms and the microgripper, the brightness is the strongest as the separation between the nanoscale parts are smaller. This confirms that MEMS and NEMS appear different in an optical microscope, especially in dark field illumination, and differences in brightness curvature has been used as a basis for image segmentation [124].

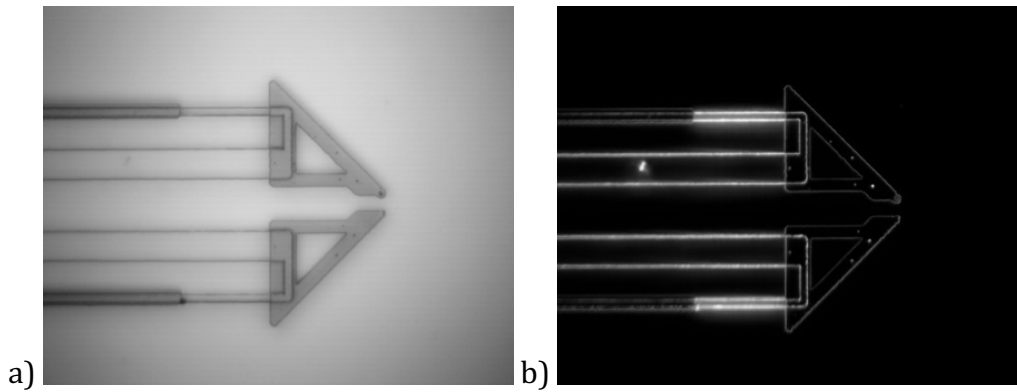


Figure 1.11: Optical images of a microscale gripper in a) bright and b) dark field illumination (courtesy Prof. R. Syms and Dr D. Liu)

Although optical imaging of MEMS and NEMS is clearly possible, the relation between the image detail and the device geometry is unclear. From the limited examples above, it is already clear that the images are formed by waves scattered from the edges of deep etched structures. The aim of this work is therefore to develop a framework for understanding the imaging process, by providing a simple method for simulating images of arbitrary device layouts using theories capable of modelling edge wave generation. The emphasis will be on devices fabricated by DRIE (since this results in deep, rectangular sections) and dark-field illumination (since this results in stronger contrast).

A key difference from earlier work is that non-periodic structures will be considered as well as periodic structures. The structures include half-planes, wedges, mesas, and adjacent mesas - the canonical objects of the geometrical theory of diffraction and its derivatives, which model scattered fields as a combination of geometric waves and edge waves [125, 126]. The first goal will be to develop methods of calculating scattered fields, including in-plane scattering between adjacent features. The second goal will be to examine the modification of the scattered field as it passes through a lens and hence predict images of 3D nanostructured objects. The remainder of the thesis is therefore organised as follows.

1.4 Thesis overview

In Chapter 2, an elementary feature — a half-plane or knife edge — is introduced as a first step towards understanding the imaging of more general device structures. Thin mesas can be easily fabricated by STL and form the basis of many elastic suspensions. Imaging

is discussed by comparing the principles of scanning electron and optical microscopy. In optical microscopy, images are formed by collecting electromagnetic (EM) waves scattered from the sample. This process is analysed in the Fourier domain in terms of complex transmission through an imaging system. Maxwell's equations are then introduced as a method of accurately calculating the scattering.

In Chapter 3, the famous Sommerfeld half-plane diffraction problem [127, 128] is reviewed together with its far-field solution, the geometrical theory of diffraction (GTD) [129-132]. The GTD uses a ray optics approach but suffers from shadow boundary and caustic problems. Solutions are obtained in terms of the Fresnel integral, which dictates the boundary singularity, and different approximations of the integral are compared. Alternative approaches — the physical theory of diffraction (PTD) [133-135] and the uniform theory of diffraction (UTD) [136-138] — are then discussed. PTD calculates the induced surface current from the incident EM wave and excites an outgoing diffracted magnetic field. The later part of the chapter discusses the theory of multiple diffraction [139-142], which can model closely spaced edges, and the incremental theory of diffraction (ITD) [143-146], which can model structures with curved or finite edges. Both are important for imaging more complex device layouts.

In Chapter 4, the geometric theory of diffraction is considered further by comparison with full-wave simulation results from FEM. Since the structures considered in GTD are isolated, isolated NEMS features are simulated using the commercial software COMSOL® Multiphysics. Spatial filtering of the back-scattered field is presented as a method of simulating the imaging process performed by a lens. The magnitudes of each backscattered diffraction order are first calculated, and the image field is obtained by recombining those diffraction orders that pass through the lens. Results obtained by the two methods are then compared, confirming agreement between numerical and analytical results for elementary device features for both TE and TM polarizations.

In Chapter 5, another diffraction theory — modal theory, which can find solutions for periodic structures — is compared with FEM simulation. This chapter also compares the difference between single and double knife edge features in terms of electric field strength. The effects of geometry are studied by comparing the results from the two

diffraction theories. Among the model parameters (feature depth, simulation domain width, feature width and number of incident waves) feature width is the only parameter that strongly impacts the electric field. Both diffraction theories describe the scattering problem almost identically and the simulated images are shown to be highly similar.

In Chapter 6, a finite element model of 3D structures is constructed and compared with the previously reviewed 2D models. Scattering at different azimuth angles is discussed and studied in both TE and TM polarizations. Dark field monochromatic images of different 3D shapes are presented and analysed. The effect of the out-of-plane illumination angles and wavelengths are also illustrated, and it is concluded that diffuse illumination of general structures can be approximated by four incident waves for structures in a Manhattan arrangement and eight waves for structures with curved edges.

The analysis methods above are all rigorous but result in extremely lengthy calculations. In Chapter 7, a novel analytic model of dark-field imaging is developed to allow much more rapid simulation of the images of general NEMS. The method simplifies complex diffraction problems by first segmenting device layouts into edges that generate scattered edge waves. The incremental theory of diffraction is used to provide diffraction coefficients for elementary sections along each edge. Each scattered component is propagated as a spherical wave, using a Huygens' type construction, and multiple scattering is computed by convolution with an appropriate kernel for in-plane radiation. Calculation time is reduced by allowing a reduced number of waves at each wavelength. Model predictions are compared with images obtained by full-wave simulation for both TE and TM polarizations, and excellent agreement is obtained despite the simplicity of the approach.

Chapter 8 summarises the arguments and conclusions of the thesis and presents some directions for possible future work.

References

MEMS

1. K. E. Petersen, "Silicon as a mechanical material", Proc. IEEE, vol. 70, pp. 420-457, 1982.
2. L. S. Fan, Y.-C. Tai, R. S. Muller, "Integrated movable micromechanical structures for sensors and actuators", IEEE Trans. Electron Dev., vol. 35, pp. 724-730, 1988.
3. K. Wise, "Integrated microelectromechanical systems: A perspective on mems in the 90s", Proc. IEEE MEMS Conf., pp. 33-38, 1991.
4. H. Seidel, M. Aikele, M. Rose, S. Toelg, "Safety relevant microsystems for automotive applications", Microsyst. Tech., vol. 7, pp. 244-248, 2002.
5. R. T. Howe, "Surface micromachining for microsensors and microactuators," J. Vac. Sci. Tech. B, vol. 6, pp. 1809-1813, 1988.
6. S. Middlehoek, S. Audet, "Silicon sensors", Academic Press, London, 1989.
7. J. Brysek, "Impact of MEMS technology on society", Sensors Actuators A, vol. 56, pp. 1-9, 1996.

MEMS applications

8. P. D. Mitcheson, T. C. Green, E. M. Yeatman, A. S. Holmes, "Architectures for vibration-driven micro-power generators", IEEE JMEMS, vol. 13, pp. 429-440, 2004.
9. S. Lucyszyn, "Review of radio frequency microelectromechanical systems (RF MEMS) technology", IEEE Proc. Sci., Meas. Tech., vol. 151, pp. 93-103, 2004.
10. R. R. A. Syms, D. F. Moore, "Optical MEMS for telecommunications", Materials Today, July/August Issue, pp. 26-35, 2002.
11. R. R. A. Syms, S. Wright, "MEMS mass spectrometers: the next wave of miniaturization", J. Micromech. Microeng., vol. 26, pp. 023001, 2016.
12. G. M. Whitesides, "The origins and future of microfluidics", Nature, vol. 442, pp. 3680373, 2006.
13. N. I. Maluf, G. T. A. Kovacs, D. Gee, "Recent advances in medical applications of MEMS", Proc. Wescon96, Anaheim, CA, 1996.
14. C. R. Grayson, R. S. Shawgo, A. M. Johnson *et al.*, "A bioMEMS review: MEMS technology for physically integrated devices", Proc. IEEE, vol. 92, pp. 6-21, 2004.

Microactuators

15. H. Fujita, "Microactuators and micromachines", Proc. IEEE, vol. 86, pp. 1721-1732,

1998.

16. E. Thielecke, E. Obermaier, "Microactuators and their technologies", *Mechatronics*, vol. 10, pp. 431-455, 2000.

NEMS

17. M. L. Roukes, "Nanomechanical systems: progress, challenges, and ultimate limits", in 58th Device Research Conf., Denver, CO, June 19-21, pp. 99-100, 2000.
18. P. N. Kambali, V. S. Nikhil, and A. K. Pandey, "Surface and nonlocal effects on response of linear and nonlinear NEMS devices", *Appl. Math. Mod.*, vol. 43, pp. 252-267, 2017.
19. D. Berman and J. Krim, "Surface science, MEMS and NEMS: Progress and opportunities for surface science research performed on, or by, microdevices", *Prog. Surf. Sci.*, vol. 88, pp. 171-211, 2013.
20. H. G. Craighead, "Nanoelectromechanical systems", *Science*, vol. 290, pp. 1532-1535, 2000.
21. K. L. Ekinci and M. L. Roukes, "Nanoelectromechanical systems", *Rev. Sci. Inst.*, vol. 76, pp. 061101, 2005.
22. S. Lyshevski and M. Lyshevski, "Systematic synthesis of nano- and micro-electromechanical systems", *Proc. 40th IEEE Conf. on Decision and Control*, vol. 1, pp. 425-430, 2001.
23. Z. Cui and C. Gu, "Nanofabrication challenges for NEMS", *Proc. 1st IEEE Int. Conf. on Nano/Micro Engng. and Mole. Sys.*, Zhuhai, China, Jan 18-21, pp. 607-610, 2006.

IC Fabrication

24. S. Sze, "VLSI fabrication", McGraw Hill, New York, 1983

MEMS Foundries

25. P. R. Salomon, "European MEMS foundries", *Proc. SPIE*, vol. 4979, pp. 327-336, 2003.
26. C. M. Liu, B. C. S. Chou, R. C.-F. Tsai *et al.*, "MEMS technology development and manufacturing in a CMOS foundry", *Proc. 16th Int. Solid State Sensors, Actuators and Microsystems Conf.*, Beijing, China, June 5-9, pp. 807-810, 2011.
27. S. Majoni, "Bosch MEMS foundry service", *Proc. ESSDERC Conf.*, Cracow, Poland, Sept 23-26, pp. 44-46, 2019.
28. D. A. Koester, R. Mahadevan, A. Shishkoff, K. W. Markus, "MUMPS Design Handbook, Rev 4", Cronos Integrated Microsystems, NC, 1999.

Processing

29. G. Delapierre, "Micromachining: a survey of the most commonly used processes", *Sensors and Actuators*, vol. 17, pp. 123-138, 1989.
30. M. J. Madou, "Fundamentals of microfabrication", CRC Press, Boca Raton, 1997.
31. J. M. Bustillo, R. T. Howe, R. S. Muller, "Surface micromachining for microelectromechanical systems", *Proc. IEEE*, vol. 86, pp. 1552-1573, 1998.

Design

32. S. D. Senturia, R. M. Harris, B. P. Johnson *et al.*, "A computer-aided design system for microelectromechanical systems (MEMCAD)", *IEEE JMEMS*, vol. 1, pp. 3-13, 1992.
33. J. V. Clarke, N. Zhou, K. S. J. Pister, "MEMS simulation using SUGAR v0.5", *Proc. Solid State Sensors and Actuators Workshop*, Hilton Head Island, SC, June 8-11, pp. 191-196, 1998.
34. J. Gilbert, A. S. Morris, "CAD tools for optical MEMS", *IEEE/LEOS International Conference on Optical MEMS*, Kauai, HI, Aug 21-24, pp. 45-46, 2000.
35. S. D. Senturia, "Microsystem design", Kluwer, Dordrecht, 2000.

Commercial software

36. <https://www.comsol.com>
37. <https://www.ansys.com>

Projection lithography

38. H. Kawata, J. M. Carter, "Optical projection lithography using lenses with numerical apertures greater than unity", *Microelectr. Engng.*, vol. 9, pp. 31-36, 1989.
39. R. H. Stulen, D. W. Sweeney, "Extreme ultraviolet lithography", *IEEE J. Quant. Elect.*, vol. 35, pp. 694-699, 1999.
40. L. R. Harriott, "Limits of lithography", *Proc. IEEE*, vol. 89, pp. 366-374, 2001.
41. M. Rothschild, "Projection optical lithography", *Materials Today*, vol. 8, pp. 18-24, 2005.
42. J. Mulken, P. C. Hinner, M. Kubise, *et al.*, "Holistic optimization architecture enabling sub-14-nm projection lithography", *J. Micro/NanoLith., MEMS and MOEMS*, vol. 13, pp. 01106, 2014.

NEMS Foundries

43. E. Myers, "Large scale integration of nanosystems", *Proc. ECS Meeting Abstracts*, MA2011-01, 2011.
44. J. L. Munoz-Gamarra, P. Alcaine, E. Marrigi *et al.*, "Integration of NEMS resonators in a 65 nm CMOS technology", *Microelectr. Engng.*, vol. 110, pp. 246-249, 2013.

Photolithography

45. W. M. Moreau, "Semiconductor lithography", Plenum Publishing Co., New York, 1988
46. G. L.-T. Chiu, J. M. Shaw, "Optical lithography — introduction", IBM J. Res. Dev., vol. 41, pp. 3-6, 1997.
47. D. G. Flagelo, B. Arnold, "Optical lithography for nanotechnology", Proc. SPIE, vol. 6327, pp. 63270D, 2006.
48. H. Lorenz, M. Despont, N. Fahrni *et al.*, "SU-8: a low-cost negative resist for MEMS", J. Micromech. Microeng., vol. 7, pp. 121-124, 1997.
49. A. R. Mirza, S. Heidemann, E. Sisson, "Resist technology for MEMS applications Micromachine Dev., vol. 5, pp. 1-3, 2000.

Nanoimprint lithography

50. S. Y. Chou, P. R. Krauss and P. J. Rentstrom, "Nanoimprint lithography", J. Vac. Sci. Tech. B, vol. 14, pp. 4129-4133, 1996.
51. A. Lebib, Y. Chen, J. Bourneix *et al.*, "Nanoimprint lithography for a large area pattern replication", Microelectr. Engng., vol. 46, pp. 319-322, 1999.
52. H. Lee, G. Y. Jung, "Full wafer scale near zero residual nano-imprinting lithography using UV curable monomer solution", Microelectr. Engng., vol. 77, pp. 42-47, 2005.

Electron beam lithography

53. Y. Chen, "Nanofabrication by electron beam lithography and its applications: a review", Microelectr. Engng., vol. 135, pp. 57-72, 2015.
54. C. Vieu, F. Carcenac, A. Pépin *et al.*, "Electron beam lithography: resolution limits and applications", Appl. Surf. Sci., vol. 164, pp. 111-117, 2000.
55. K. Yamazaki, H. Namatsu, "5 nm-order electron beam lithography for device fabrication", Jpn. J. Appl. Phys., vol. 43, pp. 3767-3771, 2004.

Scanning probe lithography

56. Z. J. Davies, G. Abadal, O. Hansen, X. Borisé, "AFM lithography of aluminium for nano-electromechanical systems", Ultramicroscopy, vol. 97, pp. 467-472, 2003.
57. R. Garcia, A. W. Knoll, E. Riedo, "Advanced scanning probe lithography", Nature Nanotechnology, vol. 9, pp. 577-587, 2014
58. P. Fan, J. Gao, H. Mao *et al.*, "Scanning probe lithography: state of the art and future perspectives", Micromachines, vol. 13, pp. 228, 2022.

Focused ion beam lithography

59. J. Meingalis, "Focused ion beam lithography", Nucl. Instrum. Meth. B, vol. 80-81, pp. 1271-1280, 1993.
60. K. Arshak, M. Mihov, A. Arshak *et al.*, "Focused ion beam lithography – overview and new approaches", Proc. 24th Int. Conf. on Microelectronics, Nis, Serbia, May 16-19, pp. 459-462, 2004.

Dip pen lithography

61. R. D. Pine, J. Zhu, F. Xu, S. Hong and C. A. Mirkin, "Dip-pen nanolithography", Science, vol. 283, pp. 5402, 1999.
62. G. Liu, M. Hirtz, H. Fuchs, Z. Zheng, "Development of dip-pen nanolithography (DPN) and its derivatives", Small, vol. 15, 1900564, 2019.

Nanosphere lithography

63. J. C. Hulteen, R. P. Van Dyne, "Nanosphere lithography: a materials general fabrication process for periodic surface arrays", J. Vac. Sci. Tech. A, vol. 13, pp. 1553-1558, 1995.
64. P. Colson, C. Henrist, R. Cloots, "Nanosphere lithography: a powerful method for the controlled manufacturing of nanomaterials", J. Nanomaterials, vol. 2013, pp. 1-19, 2013.

Island lithography

65. M. Green, T. Tsuchiya, "Mesoscopic hemisphere arrays for use as resist in structure fabrication", J. Vac. Sci. Tech. B, vol. 17, pp. 2074-2083, 1999.
66. M. Green, T. Tsuchiya, R. R. A. Syms, "Structural fabrication using cesium chloride island arrays as resist in a fluorocarbon reactive ion etching plasma", Electrochem. Solid-State Letters, vol. 3, pp. 44-46, 2000.
67. M. Green, F. M. Liu, "SERS substrates fabricated by island lithography: the silver/pyridine system", J. Phys. Chem. B, vol. 107, pp. 13015-13021, 2003.

Wet etching

68. J. Haneveld, H. Jansen, E. Berenschot *et al.*, "Wet anisotropic etching for fluidic 1D nanochannels", J. Micromech. Microeng., vol. 13, pp. S62-S66, 2003.
69. D. B. Lee, "Anisotropic etching of silicon", J. Appl. Phys., vol. 40, pp. 4569-4574, 1969.
70. G. T. A. Kovacs, N. I. Maluf, K. E. Petersen, "Bulk micromachining of silicon", Proc. IEEE, vol. 86, pp. 1536-1551, 1998.
71. Y.-I. Lee, K.-J. Park, J. Lee *et al.*, "Dry release for surface micromachining with HF vapor-phase etching", IEEE JMEMS, vol. 6, pp. 226-233, 1997.

72. D. Xu, X. Xiong, Y. Wang, "Isotropic silicon etching with XeF₂ gas for wafer-level micromachining", IEEE JMEMS, vol. 21, pp. 1436-1444, 2012.

DRIE

73. J. K. Bhardwaj, H. Ashraf, "Advanced silicon etching using high density plasmas", Proc. SPIE, vol. 2639, pp. 224-233, 1995.
74. H. Jansen, H. Gardeniers, M. de Boer *et al.*, "A survey on the reactive ion etching of silicon in microtechnology", J. Micromech. Microeng., vol. 6, pp. 14-28, 1996.
75. A. M. Hynes, H. Ashraf, J. K. Bhardwaj *et al.*, "Recent advances in silicon etching for MEMS using the ASETM process", Sensors and Actuators, vol. 74, pp. 13-17, 1999.
76. W. T. Pike, W. Karl, S. Kumar, S. Vijendran, T. Semple, "Analysis of sidewall quality in through-wafer deep reactive-ion etching", Microelectr. Engng., vol. 73-74, pp. 340-345, 2004.

CVD

77. C. Mastrangelo, Y.-C. Tai, R. S. Muller, "Thermophysical properties of low-residual stress silicon-rich LPCVD silicon nitride films", Sensors and Actuators A, vol. 21-A23, pp. 856-860, 1990.
78. A. Stoffel, A. Kovacs, W. Kronast *et al.*, "LPCVD against PECVD for micromechanical applications", J. Micromech. Microeng., vol. 6, pp. 1-13, 1996.
79. A. Tarraf, J. Daleiden S. Irmer, D. Prasai, H. Hilmer, "Stress investigation of PECVD dielectric layers for advanced optical MEMS", J. Micromech. Microeng., vol. 14, pp. 317-323, 2004.

Epitaxy

80. G. B. Stringfellow, "Epitaxy", Rep. Prog. Phys., vol. 45, pp. 469-525, 1984.
81. R. C. Anderson, R. S. Muller, C. W. Tobias, "Porous polycrystalline silicon: a new material for MEMS", IEEE JMEMS, vol. 3, pp. 10-18, 1994.
82. A. Villa, B. de Masi, A. Corigliano, C. Comi, "Mechanical characterization of epitaxial polysilicon in MEMS", Proc. Comp. Fluid Sol. Mech. Conf., June 17-20, 2003

BSOI

83. B. Diem, P. Rey, S. Renard *et al.*, "SOI 'SIMOX': from bulk to surface micromachining, a new age for silicon sensors and actuators", Sensors and Actuators A, vol. 46, pp. 8-16, 1995.
84. A. Benitez, J. Esteve, J. Bausells, "Bulk silicon microelectromechanical devices fabricated from commercial bonded and etched-back silicon-on-insulator

substrates”, *Sensors and Actuators A*, vol. 50, pp. 99-103, 1995.

85. E. H. Klaassen, K. Petersen, J. M. Noworolski *et al.*, “Silicon fusion bonding and deep reactive ion etching: a new technology for microstructures”, *Sensors and Actuators A*, vol. 52, pp. 132-139, 1996.

Deposition

86. J. Safarian, T. A. Engh, “Vacuum evaporation of pure metals”, *Metal. Mats. Trans. A*, vol. 44, pp. 747-753, 2013.
87. P. D. Davidse, “Theory and practice of RF sputtering”, *Vacuum*, vol. 17, pp. 139-145, 1967.
88. J. L. Vossen, “Control of film properties by RF sputtering techniques”, *J. Vac. Sci. Tech.*, vol. 8, pp. S12-S30, 1971.

Electroplating

89. Q. Shi, S.-C. Chang, M. W. Putty, D. B. Hicks, “Characterization of electroformed nickel microstructures”, *Proc. SPIE*, vol. 2639, pp. 191-199, 1995.
90. C. C. Tsao, E. Sachs, “Photo-electroforming: 3-D geometry and materials flexibility in a MEMS fabrication process”, *IEEE JMEMS*, vol. 8, pp. 161-171, 1999.
91. M. Wycisk, T. Tonnesen, J. Binder, S. Michaelis, H. J. Timme, “Low-cost post-CMOS integration of electroplated microstructures for inertial sensing”, *Sensors and Actuators A*, vol. 83, pp. 93-100, 2000.

HAR NEMS

92. G. Sun, X. Zhao, C. C. Kim, “Fabrication of very-high-aspect-ratio microstructures in complex patterns by photoelectrochemical etching”, *IEEE JMEMS*, vol. 21, pp. 1504–1512, 2012.
93. N. H. A. Razak, N. Amin, T. S. Kiong, K. Sopian, M. Akhtaruzzaman, “High-aspect-ratio silicon nanostructures on n-type silicon wafer using metal-assisted chemical etching (MACE) technique”, in *2021 IEEE 48th Photovoltaic Specialists Conf. (PVSC)*, Fort Lauderdale, FL, June 20-25, pp. 2596-2599, 2021.
94. P. V. Antonov, M. R. Zuiddam, J. W. Frenken, “Fabrication of high-aspect ratio silicon nanopillars for tribological experiments”, *J. Micro/Nanolithography, MEMS, and MOEMS*, vol. 14, pp. 044506, 2015.

Sidewall transfer lithography

95. W. R. Hunter, T. C. Holloway, P. K. Chatterjee, A. F. Tasch, “A new edge-defined approach for submicrometer MOSFET fabrication”, *IEEE Electron. Dev. Letts.*, vol. 2,

pp. 4-6, 1981.

96. U. Hilleringmann, T. Vierigge, J. T. Horstmann, "A structure definition technique for 25 nm lines of silicon and related material", *Microelectr. Engng.*, vol. 53, pp. 569-572, 2000.
97. Y.-K. Choi, J. Zhu, J. Grunes, et al., "Fabrication of sub-10-nm silicon nanowire arrays by size reduction lithography", *J. Phys. Chem.*, vol. 107, pp. 3340-3343, 2003.

STL NEMS

98. D. Liu and R. R. A. Syms, "NEMS by sidewall transfer lithography", *IEEE J. MEMS*, vol. 23, pp. 1366-1373, 2014.
99. D. Liu, R. R. A. Syms, M. M. Ahmad, "NEMS by multilayer sidewall transfer lithography", in *2015 28th IEEE Int. Conf. on MicroElectroMechanical Systems (MEMS)*, Estoril, Portugal, Jan 18-22, pp. 288-291, 2015.
100. R. R. A. Syms, D. Liu, M. M. Ahmad, "Nanostructured 2D cellular materials in silicon by sidewall transfer lithography NEMS", *J. Micromech. Microeng.*, vol. 27, pp. 075003, June 2017.

Simulation of lithography

101. E. C. Kintner, "Method for the calculation of partially coherent imagery", *App. Opt.*, vol. 17, pp. 2747-2753, 1978.
102. D. Nyyssonen, "Theory of optical edge detection and imaging of thick layers", *JOSA*, vol. 72, pp. 1425-1436, 1982.
103. C.J.R. Sheppard, J.M. Heaton, "Images of surface steps in coherent illumination", *Optik*, vol. 68, pp. 267-280, 1984.
104. J. T. Sheridan, C. J. R. Sheppard, "Coherent imaging of periodic thick fine isolated structures", *JOSA A*, vol. 10, pp. 614-632, 1993.
105. J. T. Sheridan, C. J. R. Sheppard, "Modelling of images of square-wave gratings and isolated edges using rigorous diffraction theory", *Opt. Comm.*, vol. 105, pp. 367-378, 1994.
106. J. T. Sheridan, T.O. Köurner, "Imaging periodic surface relief structures", *J. Microscopy*, vol. 177, pp. 95-107, 1995.

Modal theory

107. T. Tamir, H. C. Wang, A. A. Oliner, "Wave propagation in sinusoidally stratified dielectric media", *IEEE Trans. Micro. Theory Tech.*, vol. MTT-12, pp. 323-225, 1964.
108. F. G. Kaspar, "Diffraction by thick, periodically stratified gratings with complex

dielectric constant”, J. Opt. Soc. Am., vol. 63, pp. 37-45, 1973.

109. F. G. Kaspar, “Computation of light transmitted by a thick grating, for application to contact printing”, J. Opt. Soc. Am., vol. 64, pp. 1623-1630, 1974.

Rigorous coupled wave theory

110. M. G. Moharam, T. K. Gaylord, “Rigorous coupled-wave analysis of planar grating diffraction”, J. Opt. Soc. Am., vol. 71, pp. 811-18, 1981.
111. M. G. Moharam, T. K. Gaylord, “Diffraction analysis of dielectric surface-relief gratings”, J. Opt. Soc. Am., vol. 72, pp. 1385-92, 1982.

Analysis of microscopy

112. A. R. Stokes, “Three-dimensional diffraction theory of microscope image formation”, Proc. Roy. Soc. Lond. A, vol. 212, pp. 64-274, 1952.
113. H. H. Hopkins, “On the diffraction theory of optical images”, Proc. Roy. Soc. Lond. A, vol. 217, pp. 408-432, 1953
114. B. Richards, E. Wolf, “Electromagnetic diffraction in optical systems. II. Structure of the field in an aplanatic system”, Proc. Roy. Soc. A, vol. 253, pp. 358-379, 1959.
115. B.R. Frieden, “Optical transfer of a three-dimensional object”, JOSA, vol. 57, pp. 56-66, 1967.

Optical metrology of MEMS

116. W. Osten, B. J. Thompson, A. Duparre *et al.*, “Optical inspection of microsystems”, CRC Press, Boca Raton, 2007.
117. G. Dussler, B. Broecher, T. Pfeifer, “Inspection of microsystems with a laser scanning microscope”, SPIE Proc., vol. 3825, pp. 144-150, 1999.
118. A. Bosseboeuf, S. Petitgrand, “Application of microscopic interferometry techniques in the MEMS field”, Proc. SPIE, vol. 5145, pp. 1-16, 2003
119. C. Rembe, G. Siegmund, H. Steger, M. Wörtge, “Measuring MEMS in motion by laser Doppler vibrometry” in Optical inspection of MEMS, 2nd Edn., W. Osten Ed., CRC Press, Boca Raton, 2019.

Electron microscopy

120. R. F. Egerton, “Physical principles of electron microscopy: an introduction to TEM, SEM and AEM”, Springer, New York, 2016.

Optical metrology of NEMS

121. K. Creath, “Submicron linewidth measurement using an interferometric optical profiler” in Integrated Circuit Metrology, Inspection, and Process Control V, W. H.

- Arnold, Ed., vol. 1464, Int. Soc. for Optics and Photonics. SPIE, pp. 474-483, 1991.
122. S. Morgan, E. Choi, M. Somekh, C. See, "Interferometric optical microscopy of subwavelength grooves", *Optics Comms.*, vol. 187, pp. 29-38, 2001.
 123. R. R. A. Syms, A. Bouchaala, O. Sydoruk, D. Liu, "Optical imaging and image analysis for high aspect ratio NEMS", *J. Micromech. Microeng.*, vol. 29, pp. 1-11, 2018.
 124. R. R. A. Syms, A. Bouchaala, O. Sydoruk, "Improved optical imaging of high aspect ratio NEMS using dark-field microscopy", *Nanotechnology*, vol. 30, pp. 285301, 2019.

Reviews of diffraction theories

125. G. D. Durgin, "The practical behavior of various edge diffraction formulas", *IEEE Ant. Propag. Mag.*, vol. 51, pp. 24-35, 2009.
126. F. Hacivelioglu, L. Sevgi, P. Ya. Ufimtsev, "Electromagnetic scattering from wedge with perfectly reflecting boundaries: analysis of asymptotic techniques", *IEEE Ant. Propag. Mag.*, vol. 53, pp. 232-253, 2011.

Sommerfeld

127. A. J. W. Sommerfeld, "Mathematische theorie der diffraction", *Mathematische Annalen*, vol. 47, pp. 317-341, 1896.
128. A. J. W. Sommerfeld, "Optics", Academic Press, New York, 1954.

Geometrical theory of diffraction

129. J. B. Keller, "Geometrical theory of diffraction", *J.O.S.A.*, vol. 52, pp. 116-130, 1962.
130. T. B. A. Senior, P. L. E. Uslenghi, "Comparison between Keller's and Ufimtsev's theories for the strip", *IEEE Trans. Ant. Propagat.*, vol. AP-19, pp. 557-558, 1971.
131. L. Kaminetzky, J. B. Keller, "Diffraction by edges and vertices of interfaces", *SIAM J. Appl. Math.*, vol. 28, pp. 839-836, 1975.
132. P. L. Christiansen, N. Chr. Albertsen, O. Breinbjerg, "50 years with J. B. Keller's geometrical theory of diffraction in Denmark – revisiting the theory: impedance half-plane diffraction coefficients", *IEEE Ant. Propag. Mag.*, vol. 55, pp. 32-40, 2013.

Physical theory of diffraction

133. P.Ya. Ufimtsev, "Approximate computation of the diffraction of plane electromagnetic waves at certain metal bodies", *Sov. Phys. Tech. Phys.*, vol. 27, pp. 1708-1718, 1957.
134. P.Ya. Ufimtsev, "Secondary diffraction of electromagnetic waves by a strip", *Sov. Phys. Tech. Phys.*, vol. 28, pp. 535-548, 1958.

135. P. Ya. Ufimtsev, "Method of edge waves in the physical theory of diffraction", translated by U.S. Air Force Foreign Technology Division, Wright-Patterson AFB, OH, 1971.

Uniform theory of diffraction

136. R. M. Lewis, J. Boersma, "Uniform asymptotic theory of edge diffraction", J. Math. Phys., vol. 10, pp. 2291-2305, 1969.
137. R. G. Kouyoumjian, P. H. Pathak, "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface", Proc. IEEE, vol. 62, pp. 1448-1461, 1974.
138. P. H. Pathak, "Uniform geometrical theory of diffraction", Proc. SPIE, vol. 0358, 1982.

Multiple diffraction

139. J. Deygout, "Multiple knife-edge diffraction of microwaves", IEEE Trans. Ant. Propag., vol. AP-14, pp. 480-489, 1966.
140. J. D. Hunter, R. H. T. Bates, "Secondary diffraction from close edges on perfectly conducting bodies", Int. J. Electron., vol. 32, pp. 321-333, 1972.
141. C. L. Giovaneli, "An analysis of simplified solutions for multiple knife-edge diffraction", IEEE Trans. Ant. Propag., vol. AP-32, pp. 297-301, 1984.
142. R. Tiberio, R. G. Kouyoumjian, "Calculation of high-frequency diffraction by two nearby edges illuminated at grazing incidence", IEEE Trans. Ant. Propag., vol. AP-32, pp. 1186-1196, 1984.

Incremental theory of diffraction

143. K. M. Mitzner, "Incremental length diffraction coefficients", Northrop Corp. Aircraft Division Technical Report AFAL-TR-73-296, 1974.
144. R. A. Shore, A. D. Yaghjian, "Incremental diffraction coefficients for planar surfaces", IEEE Trans. Ant. Propag., vol. 36, pp. 55-70, 1988.
145. G. Pelosi, S. Maci, R. Tiberio, "Incremental length diffraction coefficients for an impedance wedge", IEEE Trans. Ant. Propag., vol. 40, pp. 1201-1210, 1992.
146. R. Tiberio, S. Maci, "An incremental theory of diffraction: scalar formulation", IEEE Trans. Ant. Propag., vol. 42, pp. 600-612, 1994.

2. Theoretical Background

This Chapter provides more detail of the theories and assumptions that underpin the work of this thesis. We begin with a brief review of electromagnetic theory — Maxwell's equations — including the wave equation, wave solutions in vectorial and scalar form, and Fourier optics. We review the operation of optical microscopes, and present an elementary theory of image formation in terms of Fourier optics. Finally, we discuss more details of nanostructures in microelectronics and MEMS, and present examples of structures for which imaging may be necessary.

2.1 Maxwell's equations

Maxwell's equations [1] represent an extension and synthesis of work by many notable scientists, including Gauss, Ampère and Faraday, and have been established for well over a century. Accounts can be found in many standard textbooks [2-5], and there has been continual emphasis on electromagnetic waves [6-8] and optical applications [9-11]. In time-dependent differential form, the equations can be written as:

$$\begin{aligned}\nabla \cdot \mathbf{D} &= \rho & \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} & \nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}\end{aligned}\quad (2.1)$$

The field quantities $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic field strength, $\mathbf{D}$ and $\mathbf{B}$ are the electric and magnetic flux density, and $\mathbf{J}$ is the current density. All are vectors, and functions of both time t and space, which may be defined in Cartesian or some other system of co-ordinates. The final term here, ρ , is the electric charge density.

The first equation is Gauss's law, and describes the relationship between the electric flux density and electric charges. The second is the magnetic field version of Gauss's law. In this case, the right hand side is zero, because magnetic monopoles are not found in nature. The third equation is Faraday's law, which describes how a time-varying magnetic flux density generates an electric field. The last equation is Ampère's law, which describes how a current generates a magnetic field. It also contains a new term unknown to Ampère, the displacement current $\partial \mathbf{D} / \partial t$. This term was key to the development of high frequency electromagnetism, which was originally added by Maxwell to represent the flow of energy between the plates of a capacitor.

Maxwell's equations are nowadays written in the compact mathematical language of vector calculus. The term ∇ is the vector differential operator 'del' or 'nabla', given in Cartesian co-ordinates by:

$$\nabla = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k} \quad (2.2)$$

Here $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit vectors in the x -, y - and z -directions. Combining the differential operator with the standard vector operations of dot and cross product leads to divergence and curl terms, written for a general vector field $\mathbf{F}$ (again, in Cartesian co-ordinates) as:

$$\begin{aligned} \nabla \cdot \mathbf{F} &= \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \\ \nabla \times \mathbf{F} &= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathbf{k} \end{aligned} \quad (2.3)$$

In addition to Maxwell's equations, there are the constitutive equations, which relate the field quantities to the different material properties of the medium of interest:

$$\mathbf{J} = \sigma \mathbf{E} \quad \mathbf{D} = \epsilon \mathbf{E} \quad \mathbf{B} = \mu \mathbf{H} \quad (2.4)$$

Here σ , ϵ and μ are the conductivity, dielectric constant and permeability, respectively. Here these are scalars, but tensors may be used to describe anisotropic materials.

An optical material is normally a dielectric with no magnetic properties (so $\mu = \mu_0$), no charges ($\rho = 0$) and zero conductivity ($\sigma = 0$). In this case the two Maxwell curl equations reduce to:

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \quad \nabla \times \mathbf{H} = \epsilon \frac{\partial \mathbf{E}}{\partial t} \quad (2.5)$$

These expressions are pleasingly symmetric. We now take the curl of both of the equations above and make use of an additional relation known as the "curl of curl" identity:

$$\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F} \quad (2.6)$$

Here $\nabla \phi$ is the gradient of a scalar function ϕ and ∇^2 is the Laplacian. Both can be written in Cartesian co-ordinates, as:

$$\begin{aligned} \nabla \phi &= \frac{\partial \phi}{\partial x} \mathbf{i} + \frac{\partial \phi}{\partial y} \mathbf{j} + \frac{\partial \phi}{\partial z} \mathbf{k} \\ \nabla^2 \mathbf{F} &= \frac{\partial^2 \mathbf{F}}{\partial x^2} + \frac{\partial^2 \mathbf{F}}{\partial y^2} + \frac{\partial^2 \mathbf{F}}{\partial z^2} \end{aligned} \quad (2.7)$$

These operations allow the curl equations to be uncoupled into time-dependent vector wave equations in $\mathbf{E}$ or $\mathbf{H}$ alone:

$$\nabla^2 \mathbf{E} = \mu_0 \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \qquad \nabla^2 \mathbf{H} = \mu_0 \varepsilon \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (2.8)$$

Of these, the equation in $\mathbf{E}$ is the most often used, since the magnetic field may always be determined from the electric field using one of the curl relations. These equations are very general, but often simpler equations are needed in complex problems.

For example, if we now assume the electric field is oscillating at a single angular frequency ω (the monochromatic light assumption), and only has a y -component (linear polarization) we may write:

$$\mathbf{E} = E_y \exp(j\omega t) \mathbf{j} \quad (2.9)$$

Here E_y is the electric field amplitude. The wave equation then reduces to the time-independent scalar equation:

$$\nabla^2 E_y + \omega^2 \mu_0 \varepsilon E_y = 0 \quad (2.10)$$

This equation is known as the Helmholtz equation. It may be written in longhand as:

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial^2 E_y}{\partial z^2} + \omega^2 \mu_0 \varepsilon E_y = 0 \quad (2.11)$$

Although a partial differential equation, this equation is still considerably easier to solve.

2.2 Electromagnetic waves

For example, wave solutions may easily be found for infinite space. Figure 2.1a shows a wave travelling in the z -direction in the (x, z) plane. In this case, there will be no change to the electric field in the x or y -direction, so that $\partial E_y / \partial x = \partial E_y / \partial y = 0$.

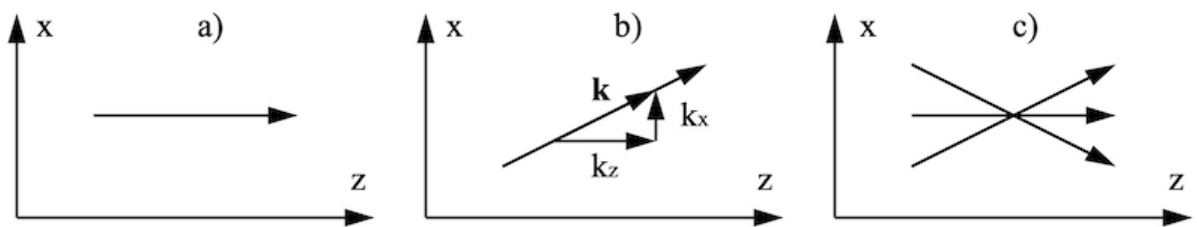


Figure 2.1: Plane wave travelling a) in the z -direction, and b) obliquely; c) multiple oblique plane waves.

The wave equation then reduces to the ordinary differential equation:

$$\frac{d^2 E_y}{dz^2} + \omega^2 \mu_0 \varepsilon E_y = 0 \quad (2.12)$$

Travelling wave solutions may be found by assuming the electric field varies as:

$$E_y(z) = E_0 \exp(-jkz) \quad (2.13)$$

Here E_0 is the wave amplitude and k is the propagation constant. Substitution into the wave equation shows that the solution is valid and yields the following relation for k :

$$k = \omega \sqrt{\mu_0 \varepsilon} \quad (2.14)$$

In free space (for which $\varepsilon = \varepsilon_0$) this yields the value $k_0 = \omega \sqrt{\mu_0 \varepsilon_0} = 2\pi/\lambda$.

The complete solution is then $\mathbf{E} = E_0 \exp\{j(\omega t - kz)\} \mathbf{j}$. This is a travelling wave, and its phase velocity may be found as:

$$v_{ph} = \omega/k = 1/\sqrt{\mu_0 \varepsilon}. \quad (2.15)$$

In free space, this yields the value $c = 1/\sqrt{\mu_0 \varepsilon_0}$, the velocity of light. In a dielectric medium, the velocity is reduced to $v_{ph} = c/\sqrt{\varepsilon_r}$ where ε_r is the relative dielectric constant. This result may be written as $v_{ph} = c/n$, where $n = \sqrt{\varepsilon_r}$ is the refractive index. This term appears extremely often in optical problems, so (for example) $k = nk_0$.

It is simple to show that the electric field is accompanied by a travelling magnetic field $\mathbf{H} = H_x \exp(j\omega t) \mathbf{i}$, with $H_x = H_0 \exp(-jkz)$. This field is travelling at the same velocity as the electric field but is polarised in the x -direction. Its amplitude may be found by making use of the Maxwell curl relations as:

$$H_0 = -\frac{k}{\omega \mu_0} E_0 \quad (2.16)$$

Thus, the two fields are both perpendicular to the direction of travel and their amplitudes are directly proportional. Solutions of this type are known as transverse electromagnetic (TEM) waves. For TEM waves, the direction of the electric field is usually defined as the polarization. In linear polarization, the electric and magnetic fields oscillate in a single direction. For circular polarization, the fields rotate clockwise or anti-clockwise in a plane perpendicular to the propagating direction. In general, the light source of a microscope is a mixture of polarizations, as an unpolarized light. For simplicity, we will only discuss two major linear polarizations in this project: transverse electric (TE) and transverse magnetic (TM) polarizations as other polarizations are the weighted sum of the two. In

particular, the TE(TM)-polarization describes the electric (magnetic) field oscillates along the y-axis (out-of-paper direction) while the wave propagates on the x-z plane. This will be further discussed in the next chapter.

In fact, the result in Equation 2.16 may be written as $H_0 = -E_0/Z_0$, where Z_0 is the impedance of the medium, given by:

$$Z_0 = \frac{\omega\mu_0}{k} = \sqrt{\frac{\mu_0}{\epsilon}} \quad (2.17)$$

Neither of these wave amplitudes are directly measurable at optical frequencies, and the power carried by the wave is generally of more interest. The instantaneous power density of an electromagnetic field is found from the Poynting vector $\mathbf{S}$:

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \quad (2.18)$$

However, because of the high oscillation frequency, it is again unusual for this expression to be made use of directly. All known optical detection systems (the eye, photographic film, photodiodes and so on) respond to the time averaged power or irradiance, given by:

$$S = \frac{1}{T} \int_0^T \mathbf{S} dt \quad (2.19)$$

The irradiance is a constant quantity. For the plane wave above, it may be found as:

$$S = \frac{1}{2Z_0} E_0^2 k \quad (2.20)$$

Thus, the wave carries power in the z-direction, and the power is proportional to the square of the electric field. More generally, it is proportional to the modulus square of the field.

Other wave solutions may easily be represented in a similar way, without needing to solve the wave equation directly. For example, a y-polarized wave travelling at an angle θ to the z-axis in the (x, z) plane as shown in Figure 2.1b may be written as:

$$E_y(x, z) = E_0 \exp\{-jk(x \sin(\theta) + z \cos(\theta))\} \quad (2.21)$$

It is simple to show by direct substitution that this solution satisfies the two-dimensional time-independent wave equation:

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2} + \omega^2 \mu_0 \epsilon E_y = 0 \quad (2.22)$$

Note that this solution may be written alternatively in terms of the components of a vector $\mathbf{k}$ (known unsurprisingly as the k-vector) as:

$$E_y(x, z) = E_0 \exp\{-j(k_x x + k_z z)\} \quad (2.23)$$

Here $k_x = k \sin(\theta)$ and $k_z = k \cos(\theta)$ are the x - and z - components of $\mathbf{k}$, whose modulus is $|\mathbf{k}| = \sqrt{k_x^2 + k_z^2}$ and whose direction represents the direction of travel.

Other common solutions to Maxwell's equations include inhomogeneous or evanescent waves (often used to represent solutions just outside dielectric waveguides), cylindrical waves (used to represent fields scattered from edges), spherical waves (often used to represent radio waves or radiation from a point), and paraxial waves (approximations to spherical waves valid near a given axis, that is often the axis of an optical instrument). Polarization is often ignored to simplify problems. In this case, spherical waves have the obvious form $E = E_0 \exp(-jkr) / r$, where the numerator accounts for phase and the denominator for amplitude, satisfying power conservation. Strictly, cylindrical waves are Bessel functions, but in the far field they tend to the simpler form $E = E_0 \exp(-jkr) / \sqrt{r}$.

Solutions to Maxwell's equations are also needed to geometries other than free space. For example, the classic problem of reflection and refraction at a boundary involves two semi-infinite media with different properties. In this case, solutions are first found for each medium as if it were infinite. The solution for the composite medium is then found by applying matching conditions at boundaries. For Maxwell's equations, the boundary conditions are inherent in the equations themselves. The full set may be written as:

$$\begin{aligned} \mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) &= 0 & \mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) &= 0 \\ \mathbf{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) &= 0 & \mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) &= \sigma_s \end{aligned} \quad (2.24)$$

Here $\mathbf{n}$ is a unit vector normal to the boundary and σ_s represents surface charge density. Of these, the two upper equations are the important ones, since they generally lead to solutions to interface problems; these require continuity of tangential electric and magnetic fields. The two lower conditions are less important, since they often lead to trivial solutions.

2.3 Fourier optics

Because Maxwell's equations are linear, more complex fields may also be constructed by superposition of known solutions. Often this is a sum of plane waves travelling in different directions as shown in Figure 2.1c. The electric field representing a set of waves with $\mathbf{k}$ -

vectors $\mathbf{k}_i$ and amplitudes e_i may be written in two dimensions as:

$$E_y(x, z) = \sum_i^N e_i \exp\{-j(k_{xi}x + k_{zi}z)\} \quad (2.25)$$

This representation is known as an angular spectrum of plane waves. Often, the waves are travelling approximately parallel to a given axis (say, the z -axis). In this case, we may make the approximation $k_{zi} \approx k$ and write:

$$E_y(x, z) = \exp(-jkz) \sum_i^N e_i \exp(-jk_{xi}x) \quad (2.26)$$

This approach is extremely powerful, since it allows general fields specified on a particular plane (say, $x = 0$) to be expanded as a sum of plane waves, as:

$$E_y(x, 0) = \sum_i^N e_i \exp(-jk_{xi}x) \quad (2.27)$$

For two dimensional fields with $\mathbf{k}$ -vectors $\mathbf{k}_{ij}$ and amplitudes e_{ij} , we clearly then have:

$$E(x, y) = \sum_i^N \sum_j^M e_{ij} \exp\{-j(k_{xi}x + k_{yj}y)\} \quad (2.28)$$

Here, for simplicity, we have assumed scalar fields at $z = 0$, written as $E(x, y)$. More generally still, the field may be expanded as an integral, written as:

$$E(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e(v_x, v_y) \exp\{-j2\pi(v_x x + v_y y)\} dv_x dv_y \quad (2.29)$$

Here $k_{x,y} = 2\pi v_{x,y}$. This equation may easily be inverted using the standard rules of Fourier transformation to find the coefficient values as:

$$e(v_x, v_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x, y) \exp\{+j2\pi(v_x x + v_y y)\} dx dy \quad (2.30)$$

The coefficients $e(v_x, v_y)$ clearly represent the spatial Fourier transform of the field $E(x, y)$. Once they are known, further propagation of the field can easily be computed, and this approach forms the basis of Fourier optics [12-15].

2.4 Optical microscopy

Optical microscopy has been used for around 400 years to visualise objects that cannot be seen with naked eye [16] and microscopes are available in many forms, depending on the wavelength and mode of illumination and the mechanism for obtaining contrast [17]. Resolution is limited by diffraction; this restricts the resolvability of the imaging to the size of a wavelength and consequently to microscale structures. Conventional microscopy is a parallel imaging process; however, other techniques such as scanning microscopy

[18, 19] build images point-by-point by scanning the object with a point source such as laser, using confocal detection to reduce the depth of focus [20]. This approach is particularly useful for biomedical objects [21] since it allows optical sectioning [22] and the formation of three-dimensional images. Similar techniques may be used for fluorescence imaging [23]. In recent decades, advanced optical “nanoscopy” techniques like stimulated emission depletion (STED) or stochastic optical reconstruction microscopy (STORM), based on super-resolved far-field fluorescence optical microscopy, have been developed [24-26]. Serial techniques such as near-field scanning optical microscopy (NSOM) have also been used to achieve a strongly sub-wavelength spatial resolution of $< 10\text{nm}$ [27, 28]. We shall not discuss these specialised techniques further. In conventional microscopy, visible light from an incoherent source, such as heated tungsten filament, is used for illumination. Other modes like ultraviolet or fluorescence microscopy use different regions of the spectrum or wavelength conversion for imaging.

Figure 2.2 shows Köhler illumination [29], which yields spatially uniform illumination regardless of the inhomogeneity of the light source. The source on the left is imaged on the aperture stop by an auxiliary lens. This image acts as the effective light source. Assuming the size of the aperture stop is much larger than the Airy disc of the auxiliary lens, the effective source can be considered incoherent. L_1 and L_2 are therefore two mutually incoherent point sources on the light source that emanate two pairs of rays, shown as solid and dotted lines. As the aperture stop lies on the focal plane, every effective point source creates a collimated beam that illuminates the object at different angles.

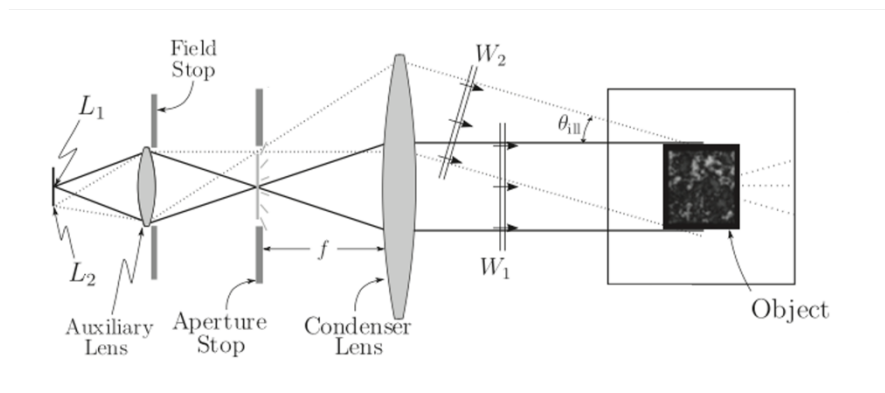


Figure 2.2: Köhler illumination [29].

Different illumination techniques are used to obtain contrast from a sample. These include bright field illumination, dark field illumination, polarization contrast illumination and phase contrast illumination. Figure 2.3 shows the schematic of an conventional light microscope in the two basic illumination modes: a) bright field (BF) and b) dark field (DF) [30]. The illumination mode is selected using an arrangement of prisms controlled by a slider bar. In BF illumination, the incident light wave from the light source (1) is reflected by a beam splitter (2), and is focused by an objective lens (3) on the sample (4). Back-scattered and reflected light is collected by the objective, passes through the beam splitter again, and is converged by an eyepiece lens (5). For DF illumination, the incident light is converted into a focused annular illumination beam by a system of mirrors (2-4). Back-scattered light from the sample (5) is then collected by the objective and is refocused by the eyepiece.

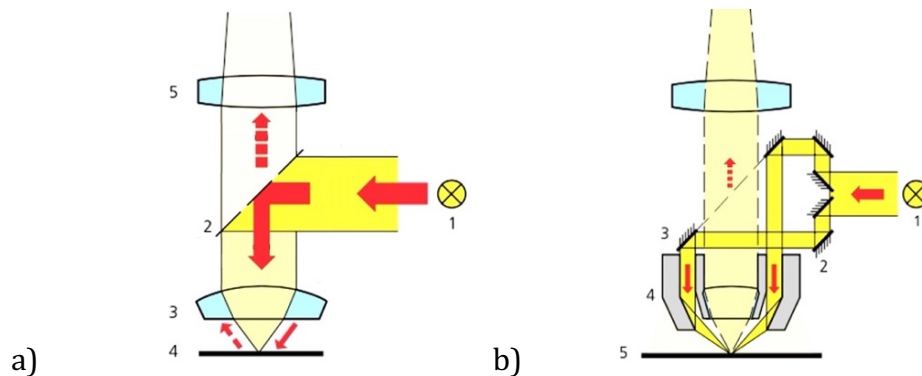


Figure 2.3: Optical microscope in a) bright and b) dark field illumination [30].

The difference between BF and DF lies in the fields collected by the lens. BF images are generated from both the diffracted and reflected fields. The reflected field is often much larger than the diffracted field, so raised features are often weakly visible. In DF illumination, the incident angles are larger. Consequently, only diffracted light within the acceptance cone of the objective is collected, and this yields much higher contrast. However, because of the weak intensity of diffracted field, DF usually requires stronger illumination. In general, DF images have larger contrast and higher resolution than BF images with the same light source. The best resolution achievable using a conventional optical microscope is typically $0.2\mu\text{m}$ with magnification below 2000x, about a thousand times smaller the resolution obtained using a scanning electron microscope.

2.5 Image formation

The theory of optical microscopy was placed on a firm foundation using Fourier optics by Hopkins [31], who treated the microscope as a linear system with the input and output as the object and image respectively. Since then, it has been extended by many others, to include partial coherence [32], three-dimensional objects [33, 34], high numerical apertures [35] and detailed simulations including vectorial effects [36, 37].

Generally, the action of the lens can be found from a combination of diffraction theory and transfer function theory, which lead to the concept of an optical transfer function [38, 39]. However, since our emphasis is on the objects themselves, a simpler explanation can be used. For example, Figure 2.4 shows a perfect unity magnification system where a single lens redirects an angular spectrum of waves diverging from a point in the object plane so that they converge to a point in the image plane. In addition to redirecting the waves, the lens restricts their angular spectrum using a pupil (which may be an aperture or the lens itself).

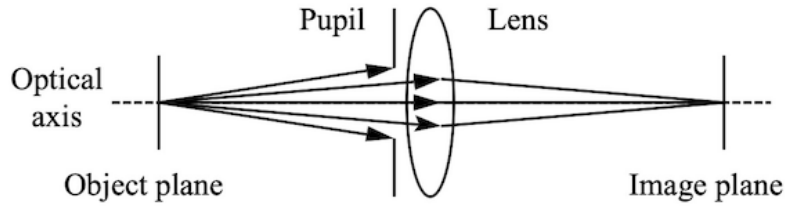


Figure 2.4: A simple optical imaging system.

Here we assume that the main effect of the lens is described by the pupil function, which we write as $P(v_x, v_y)$. Typically, an objective lens will accept radiation within a finite acceptance cone, defined by a half-angle θ_{max} . From the k-vector definition in Figure 2.1b, we can write:

$$\frac{k_{max}}{k} = \frac{2\pi v_{max}}{k} = \sin(\theta_{max}) \quad (2.31)$$

The term $\sin(\theta_{max})$ is known as the numerical aperture (NA) for a lens operating in air; for a lens operating in a different medium (for example, an immersion oil) of refractive index n , this value increases to $NA = n \sin(\theta_{max})$. Inverting the above, $v_{max} = NA/\lambda$. If we now write $v = \sqrt{v_x^2 + v_y^2}$, the pupil function is a low-pass filter:

$$P(v_x, v_y) = 1, v \leq v_{max} \quad P(v_x, v_y) = 0, v > v_{max} \quad (2.32)$$

An additional aplanatic factor $A(v_x, v_y)$ is required to ensure power conservation. This factor varies as $\sqrt{\cos(\theta)}$, where $\sin(\theta) = 2\pi v/k$. Because of the slow variation of the cosine, the aplanatic factor may often be taken as unity.

With these definitions, it can be shown that the image field obtained in a unity magnification system is:

$$E'(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(v_x, v_y) P(v_x, v_y) e(v_x, v_y) \exp\{-j2\pi(v_x x + v_y y)\} dv_x dv_y \quad (2.33)$$

Here, $e(v_x, v_y)$ represent the wave amplitudes found earlier. Thus, a coherent image may be found by Fourier transforming the object field, multiplying by the pupil and aplanatic functions, and inverse Fourier transforming the result. Both steps may be carried out very rapidly using the fast Fourier transform. Shift operations are also required to re-order the transform results before the spatial filtering operations are carried out. For diffuse or polychromatic illumination, the image may be found as the incoherent sum of a set of separate coherent object fields, obtained by illumination of the object with a set of different waves. Apart from the simulations of lithography previously mentioned [40-44], and of imaging other 3D microstructures [45-47], little attention has been paid to the calculation of the object field $E(x, y)$ or its transform $e(v_x, v_y)$ when there is multiple scattering within the object, the topic of this thesis.

2.6 Resolution

The case of a single point object is of particular interest since the size of the resulting image determines the overall resolution in general imaging. In this case, the transform e is unity. Neglecting the aplanatic function for simplicity, the image is then given by:

$$E'(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(v_x, v_y) \exp\{-j2\pi(v_x x + v_y y)\} dv_x dv_y \quad (2.34)$$

If we now adopt a polar coordinate system to write $x = r \cos(\theta)$, $y = r \sin(\theta)$, $v_x = v \cos(\phi)$ and $v_y = v \sin(\phi)$ and substitute for the pupil function, we obtain:

$$E'(x, y) = \int_0^{v_{max}} \int_0^{2\pi} \exp\{-j2\pi r v \cos(\phi - \theta)\} d\phi v dv \quad (2.35)$$

Since the aperture and image are both circularly symmetric, we can put $\theta = 0$ without loss of generality. The integral then becomes:

$$E'(r) = \int_0^{v_{max}} \int_0^{2\pi} \exp\{-j2\pi r v \cos(\phi)\} d\phi v dv \quad (2.36)$$

Integrating a first time, we obtain:

$$\int_0^{2\pi} \exp\{-j2\pi r v \cos(\phi)\} d\phi = 2\pi J_0(2\pi r v) \quad (2.37)$$

Here $J_0(x)$ is the zeroth order Bessel function of the first kind. Integrating a second time, we get:

$$E'(r) = \int_0^{v_{max}} 2\pi J_0(2\pi r v) v dv = \frac{v_{max} J_1(2\pi r v_{max})}{r} \quad (2.38)$$

Here $J_1(x)$ is the first order Bessel function. Finally, substituting for v_{max} , we obtain:

$$E'(r) = \frac{\sin(\theta_{max}) J_1(2\pi r \sin(\theta_{max})/\lambda)}{\lambda r} \quad (2.39)$$

The intensity distribution $I(r) = |E'|^2$ is then commonly written in standard form as:

$$I(r) = I_0 \left\{ \frac{2J_1(2\pi r \sin(\theta_{max})/\lambda)}{(2\pi r \sin(\theta_{max})/\lambda)} \right\}^2 \quad (2.40)$$

Here I_0 is the peak intensity. This pattern is known the point spread function (PSF) and for this simple example is an Airy disc [48]. More recent authors have calculated point images for lenses of high numerical aperture [49]. Figure 2.5a shows resulting intensity distribution, for $\lambda = 0.5\mu m$ (corresponding to green light) and $\theta_{max} = 30^\circ$. The pattern consists of a main central peak surrounded by weak concentric rings. The variation depends primarily on NA and λ ; the larger the former and the smaller the latter, the narrower the peak.

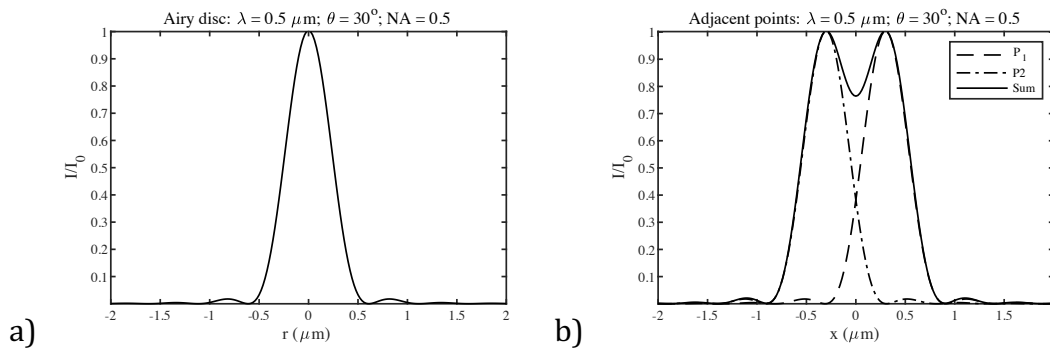


Figure 2.5: Intensity distribution for a) Airy disc, and b) images of adjacent points.

Extended images may be considered as being built up as an incoherent sum of point spread functions, one for each point in the object. Figure 2.5b shows similar patterns from two adjacent object points at a separation of $d = 0.6\mu m$, together with their sum, which

consists of two peaks separated by a valley. The two Airy discs clearly overlap, and, as the separation reduces, they will eventually become indistinguishable. This behaviour determines the resolution of the imaging system. Different definitions are commonly used for resolution, namely the Rayleigh, Abbe, and Sparrow limits [50-52], based on a minimum allowable valley depth in the sum distribution in Figure 2.5b. Table 2.1 defines these criteria and compares the values calculated using the three approaches for $\lambda = 0.5\mu\text{m}$ and $\theta_{\text{max}} = 30^\circ$ (or $NA = 0.5$). All three suggest a resolution comparable to the wavelength.

Rayleigh	Abbe	Sparrow
$\frac{1.22\lambda}{2NA}$	$\frac{\lambda}{2NA}$	$\frac{0.94\lambda}{2NA}$
$0.61\mu\text{m}$	$0.5\mu\text{m}$	$0.47\mu\text{m}$

Table 2.1: Estimates of minimum resolution d obtained using different criteria.

2.7 HAR MEMS and NEMS objects

We now consider objects likely to be encountered during simulation of NEMS imaging, based on an assessment of fabrication methods and common device features. The most flexible method of parallel nanoscale patterning is sidewall transfer lithography, alternatively known as spacer lithography and edge lithography, and we now consider its properties once again. STL was originally developed to provide a method for reducing gate length in FETs, since it determines their maximum operating frequency [53-55]. It was simple to form arrays of similar devices, but little attention was paid to integration into circuits. Subsequently, STL processes were used to form nanowires, which are predominantly needed in arrays for applications including catalysis [56-58].

Processes were developed for formation of nanoscale mesas [59, 60] (Figure 2.6a) and trenches [61, 62] (Figure 2.6b). Both were based on vertical edges formed in layers 20-100 nm thick, deposited conformally over previously etched mesas using a variety of materials including polysilicon, SiO_2 and Si_3N_4 and subsequently etched to remove all horizontal layers. The main limitations to minimum and maximum sidewall mask width are transfer of the pattern into the underlying substrate by etching (which tends to erode thin features) and stress in deposited layers, (which causes delamination of thick layers).

In addition, it remains extremely difficult to form both feature types together, for example as a nanoscale mesa adjacent to a nanoscale trench. It is also difficult to form closely spaced nanoscale feature of either type (Figure 2.6c) since this would require sub-micron resolution in the primary lithography. NEMS features are therefore likely to have nanoscale widths w spaced at microscale intervals s . Iterative methods have been developed to form secondary sidewall structures around primary sidewalls, yielding a nested array [63, 64] (Figure 2.6d). However, these arrays must again contain microscale intervals. One obvious conclusion is that many STL-fabricated structures will be visible in an optical imaging system with comparable resolution to that of the lithography system, because of the presence of microscale intervals. However, limited resolution will alter the appearance of the nanoscale parts.

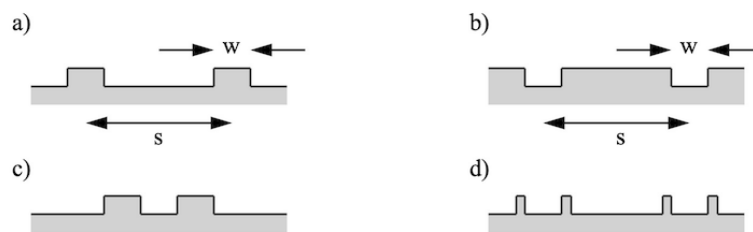


Figure 2.6: STL features: a) mesas, b) trenches, c) closely spaced mesas, d) iterated mesas.

NEMS require further adaptation of STL processing to form arrangements more complex than sets of parallel features. Assuming the devices contain movable mechanical parts, there will be a difference between nanoscale movable parts and their anchors, which need only be microscale. Other features such as electrode gaps in electrostatic drives may also be nanoscale. Thus, unlike microelectronics, typical NEMS are likely to contain a mix of dimensions. A wide variety of devices with very different operating principles have been fabricated. Examples include field emitters [65], through-wafer interconnects [66], accelerometers [67], positioning stages [68], nano-nail arrays [69], nanochannels [70, 71] and quantum dots and wires [72, 73]. There has been particular interest in fabrication of templates for nanoimprint lithography [74-77].

Here we focus on devices fabricated using a technology for NEMS fabrication developed at Imperial College, which forms anchors using conventional lithography and nanoscale suspensions using STL (Figure 2.7a) [78]. In some processes, legacy features are also formed at intersections between feature types. Multi-stage STL processes have also been

developed to allow intersecting nanoscale features [79], particularly to fabricate nanoscale cellular materials [80] (Figure 2.7b). Devices layouts are likely to be arbitrary; however, many devices are likely to follow Manhattan geometries to simplify layout. Canonical objects for simulation are therefore likely to include combinations of micro- and nano-scale features and intersecting nanoscale lines.

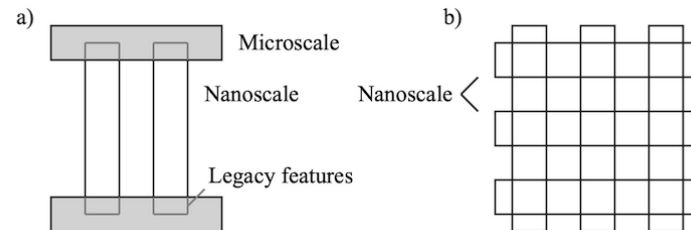


Figure 2.7: STL NEMS: a) Nanoscale suspension with microscale anchors; b) nanoscale cellular material.

The use of DRIE for pattern transfer implies that feature depths are likely to be much larger for NEMS than for microelectronics. However, the need to avoid damage to the structure by lateral erosion will limit the achievable feature depth and width. Figure 2.8 shows a representative MEMS/NEMS device, highlighting the high aspect ratio, the difference in feature width, and legacy features.

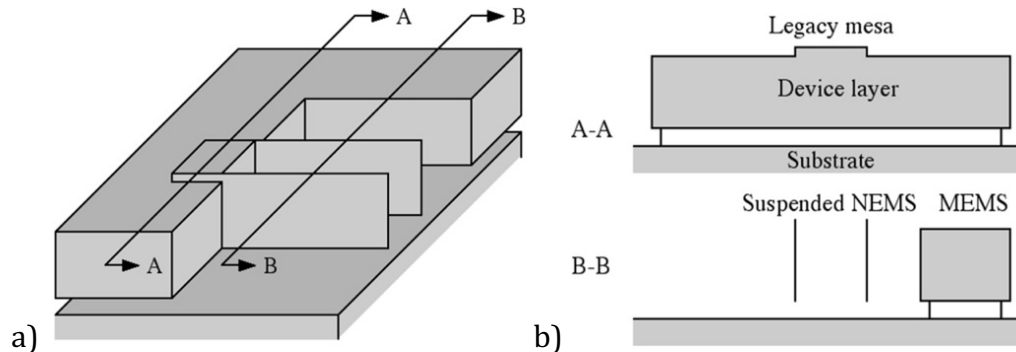


Figure 2.8: NEMS device in a) 3D and b) section (Courtesy Prof. R. Syms).

The NEMS parts consist of a set of parallel nanoscale suspension beams that are undercut and therefore free to move. HAR processing is used to force in-plane motion through a large difference in second moment of area of the two beam bending axes. The MEMS parts consist of a set of anchors rigidly attached to the substrate, which provide encasté supports for the nanoscale beam ends. Such devices can form the basis on resonant physical sensors with in-plane motion. The large feature depth implies that effect of the substrate during imaging is likely to be limited, since it will be out-of-focus. This aspect may be exploited to simplify further a model of object scattering.

2.8 Summary

An elementary account of Maxwell's equations has been presented, concentrating on those aspects important to this work, namely wave equations, electromagnetic waves, Fourier optics and imaging. Objects likely to be encountered during optical imaging of NEMS have also been discussed. In the next chapter, the geometrical theory of diffraction will be reviewed, as a possible method of simulating scattering from device structures of this type. For example, a nanoscale rib may be modelled as a half-plane, while each edge of a microscale mesa or trench corresponds to a right-angle wedge.

References

Electromagnetic theory

1. J. C. Maxwell, "A treatise on electricity and magnetism", Clarendon Press, Oxford, 1873.
2. J. A. Stratton, "Electromagnetic theory", McGraw Hill Book Company, New York, 1941.
3. E. Weber, "Electromagnetic fields", John Wiley and Sons, New York, 1950.
4. P. C. Clemmow, "Electromagnetic theory", Cambridge University Press, Cambridge, 1973.
5. L. Solymar, "Lectures on electromagnetic theory", Oxford University Press, Oxford, 1976.

Electromagnetic waves

6. Lord Rayleigh F. R. S., "On the electromagnetic theory of light", The London, and Dublin Phil. Mag., vol. 12, pp. 81-101, 1881.
7. J. A. Stratton, L. J. Chu., "Diffraction theory of electromagnetic waves", Phys. Rev. vol. 56, pp. 99-107, 1939.
8. J. A. Kong, "Electromagnetic waves", Wiley Interscience, New York, 1975.

Optics

9. M. Born, E. Wolf, "Principles of optics: electromagnetic theory of propagation, interference and diffraction of light", Pergamon Press, Oxford, 1959.
10. S. G. Lipson, H. Lipson, "Optical physics", Cambridge University Press, Cambridge, 1981.
11. R.D. Guenther, "Modern optics", John Wiley and Sons, New York, 1990.

Fourier optics

12. P. Elias, D. S. Grey, D. Z. Robinson, "Fourier treatment of optical processes", JOSA, vol. 42, pp. 127-134, 1952.
13. E. G. Steward, "Fourier optics: an introduction", Ellis Horwood, Chichester, 1983.
14. J. W. Goodman, "Introduction to Fourier optics", McGraw-Hill, New York, 1994
15. E. Huggins, "Introduction to Fourier optics", The Physics Teacher, vol. 45, pp. 364-368, 2007.

Optical Microscopy

16. M. W. Davidson, M. Abramowitz, "Optical microscopy", Encyclopedia of Imaging Science and Technology, vol. 2, pp. 1106-1141, Wiley, New York, 2002.

17. C. J. R. Sheppard, "The optics of microscopy", J. Opt. A.: Pure Appl. Opt., vol. 9, pp. S1–S6, 2007.

Scanning microscopy

18. T. Wilson, "Theory and practice of scanning optical microscopy", Academic Press, London, 1984.
19. T. Wilson, "Scanning optical microscopy", Scanning, vol. 7, pp. 61-108, 1985.
20. C. J. R. Sheppard, T. Wilson, "Image formation in confocal scanning microscopes", Optik, vol. 55, pp. 331-342, 1980.
21. T. Wilson, "Confocal microscopy", Ch. 1 in "Biomedical optical imaging", J. Fujimoto and D. L. Farkas, Eds., Oxford University Press, Oxford, 2009.
22. T. Wilson, "Resolution and optical sectioning in the confocal microscope", J. Microscopy, vol. 244, pp. 113-121, 2011.
23. T. Wilson, "Optical sectioning in confocal fluorescence microscopes", J. Microscopy, vol. 154, pp. 143-156, 1989.

Alternative methods of microscopy

24. H. Blom, J. Widengren, "Stimulated emission depletion microscopy", Chem. Rev., vol. 117, pp. 7377-7427, 2017.
25. M.J. Rist, M. Bates, X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)", Nature Meth., vol. 3, pp. 793-796, 2003.
26. A. Diaspro, P. Bianchini, "Optical nanoscopy", Riv. Nuovo Cim., vol. 43, pp. 385-455, 2020.
27. S. Kawata, Y. Inouye, "Scanning probe optical microscopy using a metallic probe tip", Ultramicroscopy, vol. 57, pp. 313-317, 1995.
28. T. Tachizaki, T. Nakata, K. Zhang, I. Yamakawa, S.-I. Taniguchi, "Nanometer-precise optical length measurement using near-field scanning optical microscopy with sharpened single carbon nanotube probe", Ultramicroscopy, vol. 186, pp. 18-22, 2018.

Microscope operation

29. A. Köhler, "New method of illumination for photomicrographical purposes", J. Royal Microscopical Soc, vol. 10, pp. 433-440, 1894.
30. F. K. Möllring, "Microscopy from the very beginning", C. Zeiss, Oberkochen, West Germany, 1967.

Theory of image formation

31. H. H. Hopkins, N. F. Mott, "On the diffraction theory of optical images," *Proc. Roy. Soc. Lond. Ser. A*, vol. 217, pp. 408–432, 1953.
32. E. C. Kintner, "Method for the calculation of partially coherent imagery", *App. Opt.* vol. 17, pp. 2747–2753, 1978.
33. E. Evans, "Comparison of the diffraction theory of image formation with the three-dimensional, first Born scattering approximation in lens systems", *Opt. Comm.*, vol. 2, pp. 317-320, 1970.
34. C. J. R. Sheppard and X. Q. Mao, "Three-dimensional imaging in a microscope", *JOSA* vol. 6, pp. 1260-1269, 1989.
35. C. J. R. Sheppard and H. J. Mathews, "Imaging in high-aperture optical systems", *JOSA A*, vol. 4, pp. 1354-1360, 1987.
36. A. Erhardt, G. Zinser, D. Komitowski, J. Bille, "Reconstructing 3-D light microscopic images by digital image processing", *Appl. Opt.*, vol. 24, pp. 194-200, 1985.
37. R. Capoglu, J. D. Rogers, A. Taflove, V. Backman, "The microscope in a computer: image synthesis from three-dimensional full-vector solutions of Maxwell's equations at the nanometer scale", *Prog. Opt.*, vol. 57, pp. 1-91, 2012.

Optical transfer functions

38. F. D. Smith, "Optical image evaluation and the transfer function", *Appl. Opt.*, vol. 2, pp. 335-350, 1963.
39. K. Rosenhauer, K. J. Rosenbruch, "The measurement of the optical transfer functions of lenses", *Rep. Progr. Phys.*, vol. 30, pp. 1–25, 1967.

Simulation of lithography

40. D. Nyyssonen, "Theory of optical edge detection and imaging of thick layers", *JOSA*, vol. 72, pp. 1425–1436, 1982.
41. C. J. R. Sheppard, J. M. Heaton, "Images of surface steps in coherent illumination", *Optik*, vol. 68, pp. 267-280, 1984.
42. J. T. Sheridan, C. J. R. Sheppard, "Coherent imaging of periodic thick fine isolated structures", *JOSA A*, vol. 10, pp. 614-632, 1993.
43. J. T. Sheridan, C. J. R. Sheppard, "Modelling of images of square-wave gratings and isolated edges using rigorous diffraction theory", *Opt. Comm.*, vol. 105, pp. 367-378, 1994.
44. J. T. Sheridan, T. O. Köurner, "Imaging periodic surface relief structures", *J. Microscopy*, vol. 177, pp. 95-107, 1995.

Simulation of general microstructure imaging

- 45. G. L. Wojcik, D. K. Vaughan, L. K. Galbraith, "Calculation of light scatter from structures on silicon surfaces", *Proc. SPIE*, vol. 774, pp. 21-32, 1987.
- 46. J. Huttunen, J., Turunen, "Phase images of grooves in a perfectly conducting surface", *Opt. Comm.*, vol. 119, pp. 485-490, 1995.
- 47. S. van Haver, J. J. M. Braat, A. J. E. M. Janssen *et al*, "Vectorial aerial-image computations of three-dimensional objects based on the extended Nijboer-Zernike theory", *JOSA A*, vol. 26, pp. 1221-1234, 2009.

Airy disc

- 48. G. B. Airy, "On the diffraction of an object glass with circular aperture", *Trans. Camb. Phil. Soc.*, vol. 5, pp. 283-291, 1835.
- 49. C. J. R. Sheppard, T. Wilson, "The image of a single point in microscopes of large numerical aperture", *Proc. Roy. Soc. Lond. A*, vol. 379, pp. 145-158, 1982.

Resolution

- 50. Lord Rayleigh F. R. S. , "Investigations in optics, with special reference to the spectroscope", *Phil. Mag.*, vol. 5, pp. 261-274, 1879.
- 51. H. Volkmann, "Ernst Abbe and his work", *Appl. Opt.*, vol. 5, pp. 1720-1731, 1966.
- 52. C. M. Sparrow, "On spectroscopic resolving power", *Astrophys. J.*, vol. 44, pp. 76-86, 1918.

Transistors

- 53. W. R. Hunter, T. C. Holloway, P. K. Chatterjee, A. F. Tasch, "A new edge-defined approach for submicrometer MOSFET fabrication", *IEEE Electron. Dev. Letts.*, vol. 2, pp. 4-6, 1981.
- 54. D. H. Kim, S.-K. Sung, J. S. Sim, K. R. Kim, J. D. Lee, B. G. Park, B. H. Choi, S. W. Hwang, D. Ahn, "Single-electron transistor based on a silicon-on-insulator quantum wire fabricated by a side-wall patterning method", *Appl. Phys. Lett.* vol. 79, pp. 3812-3814, 2001.
- 55. Y.-K. Choi, T.-J. King, C. Hu, "A spacer patterning technology for nanoscale CMOS", *IEEE Trans. Electron Devices*, vol. 49, pp. 436-441, 2002.

Nanowires

- 56. Y.-K. Choi, J. Zhu, J. Grunes, J. Bokor, G. A. Somorjai, "Fabrication of sub-10-nm silicon nanowire arrays by size reduction lithography", *J. Phys. Chem.*, vol. 107, pp. 3340-3343, 2003.

57. X. M. Yan, S. Kwon, A. M. Contreras, M. M. Koebel, J. Bokor, G. A. Somorjai, "Fabrication of dense arrays of platinum nanowires on silica, alumina, zirconia and ceria surfaces as 2-D model catalysis", *Catalysis Letts.*, vol. 105, pp. 127-132, Dec. 2005.
58. G. Kostovski, A. Mitchell, A. Holland, M. Austin, "Sidewall corrugation lithography: bulk fabrication of ordered nanowires, nanoribbons and nanorings", *Appl. Phys. Lett.*, vol. 92, pp. 22319, 2008.

STL ridges

59. E.P. Harris, R.W. Keyes, "Method for fabricating ultranarrow metallic lines", US patent 4093503, 1978.
60. U. Hilleringmann, T. Vierigge, J. T. Horstmann, "A structure definition technique for 25 nm lines of silicon and related material", *Microelectr. Engng.*, vol. 53, pp. 569-572, 2000.

STL slots

61. I. T. Ho, J. Riseman, "Method for forming a narrow dimensioned mask opening on a silicon body utilizing reactive ion etching", US patent 4209349, June 1980.
62. J. Sakimoto, H. Kawata, M. Yasuda, H. Yoshiko, "25 nm wide silicon trench fabrication by edge lithography", *Jpn. J. Appl. Phys.*, vol. 50, pp. 08KC03, Aug. 2011.

Iterated STL

63. K.-H. Chung, W. Y. Choi, S.-K. Sung, D. H. Kim, J. D. Lee, B. G. Park, "Pattern multiplication method and the uniformity of nanoscale multiple lines", *J. Vac. Sci. Tech. B*, vol. 21, pp. 1491-1495, 2003.
64. A. Carlson, T.-J. K. Liu, "Negative and iterated spacer lithography process for low variability and ultra-dense integration", *SPIE Proc.*, vol. 6924, pp. 69240B, 2008.

NEMS

65. D. S. Y. Hsu, H. F. Gray, "Vertical thin-film-edge field emitters: fabrication by chemical beam deposition, imaging of cathodoluminescence and characterisation of emission", *Thin Solid Films*, vol. 286, pp. 92-97, 1996.
66. C. Liu, "Through-wafer electrical interconnects by sidewall photolithographic patterning", *Proc. IEEE Instrum. Meas. Tech. Conf.*, St. Paul, MN, May 19-21, pp. 1402-1405, 1998.
67. X. Li, B. Cheng, Y. Wang, L. Gu, "A trench-sidewall single-wafer-MEMS technology and its typical application in high-performance accelerometers", *Proc. IEDM Tech. Dig.*, San Francisco, CA, Dec 13-15, pp. 43-46, 2004.

68. L. Gu, X. Li, H. Bao, B. Liu *et al.*, "Single-wafer processed nano-positioning XY-stages with trench-sidewall micromachining technology" J. Micromech. Microeng. vol. 16, pp. 1349-1357, 2006.
69. X. Han, H. Luo, W. G. Wu, J. Xu, G. Yan, Y. F. Jin and Y. L. Hao, "Stacked SiO₂/Si nanonail array fabricated by spacer technology for biomedical applications", Proc. Transducers 2007, Lyon, France, Jun. 10-14, pp. 1609-1612, 2007.
70. Q. Xie, Q. Zhou, F. Xie, J. Sang, W. Wang, H. A. Zhang, W. Wu, Z. Li, "Wafer-scale fabrication of high-aspect ratio nanochannels based on edge-lithography technique", Biomicrofluidics, vol. 6, pp. 016502-016502-08, 2012.
71. E. Cheng, H. Zou, X. Yin, P. Jurcicek, X. Zhang, "Fabrication of 2D polymer nanochannels by sidewall lithography and hot embossing", J. Micromech. Microeng. vol. 23, pp. 075022, 2013.
72. J. C. Yi, "Miniband properties of superlattice quantum dot arrays fabricated by the edge-defined nanowires", Microelectronics. J., vol. 39, pp. 369-374, 2008.
73. F. M. Gomez-Campos, S. Rodriguez-Bolivar, A. Luque-Rodriguez, J. A. Lopez-Villanueva, "Intraband photon absorption in edge-defined nanowire superlattices for optoelectronic applications", J. Appl. Phys., vol. 108, pp. 12307, 2010.

Nanoimprint templates

74. P. B. Grabiec, M. Zaborowski, K. Domanski, T. Gotszalk, I. W. Rangelow, "Nano-width lines using lateral pattern definition technique for nanoimprint template fabrication", Microelectr. Engng., vol. 73-74, pp. 599-603, 2004.
75. J. Haneveld, E. Berenschot, P. Maury, H. Jansen, "Nano-ridge fabrication by local oxidation of silicon edges with silicon nitride as a mask", J. Micromech. Microeng., vol. 16, pp. S24-S28, 2006.
76. Y. Zhao, E. Berenschot, M. de Boer, H. Jansen, N. Tas, J. Huskens, M. Elewenspoek, "Fabrication of a silicon oxide stamp by edge lithography reinforced with silicon nitride for nanoimprint lithography", J. Micromech. Microeng., vol. 18, pp. 064013, 2008.
77. J. Rao, H. Zou, R. R. A. Syms, E. Cheng, C. Liu, "Fabrication of 2D silicon nano-mold based on sidewall transfer", Micro and Nano Letts., vol. 6, pp. 29-33, 2011.

STL NEMS

78. D. Liu, R. R. A. Syms, "NEMS by sidewall transfer lithography", IEEE J. Microelectromech. Syst., vol. 23, pp. 1366-1373, 2014.

79. D. Liu, R. R. A. Syms, "NEMS by multilayer sidewall transfer lithography", 28th IEEE MEMS. Conf., Estoril, Portugal, Jan. 18-22, pp. 288-291, 2015.
80. R. R. A. Syms, D. Liu, M. M. Ahmad, "Nanostructured 2D cellular materials by sidewall transfer lithography NEMS", J. Micromech. Microeng., vol. 27, pp. 075003, 2017.

3. Diffraction Theory of Isolated Obstacles

This section presents several high-frequency asymptotic methods for solving diffraction problems, particularly Sommerfeld's half-plane problem and the related wedge problem. We start by considering the background and demonstrate the derivation of Sommerfeld's solution. Due to the complexity of his solution, different asymptotic approximations have been developed. We therefore review several such theories: the geometrical theory of diffraction (GTD), the uniform theory of diffraction (UTD), the physical theory of diffraction (PTD) and the incremental theory of diffraction (ITD). Finally, to connect diffraction theory to optical imaging of NEMS devices, we discuss multiple diffraction.

3.1 Sommerfeld's half-plane diffraction problem

Wave diffraction occurs when a light wave is incident on an obstacle. For example, a cylindrical wave, known as edge wave, is generated at any corner. This diffraction phenomenon was first studied by Young in early 19th century [1]. Further investigations were then carried out by Fresnel, Kirchhoff, and Helmholtz [2-4], and diffraction of EM waves by a half-plane was rapidly established as a popular canonical problem. In late 19th century, Sommerfeld studied the diffraction from a perfectly conducting thin plate with semi-infinite length and depth, famously known as the half-plane problem, and proposed the first exact solution [5, 6]. Later authors have considered diffraction from wedges, of which the half-plane represents a degenerate case. Because of the feature sizes involved, both are relevant to optical imaging and metrology of NEMS.

For example, the scattering from a HAR nanostructure is analogous to the diffraction from a half-plane (Figure 3.1a). The incident rays (black arrows) are reflected from the flat surface as retro-reflected rays (green arrows). When the incident ray is diffracted at the tip of the half-plane, the diffracted wave (blue) is radiated in all directions with angle-dependent amplitude. Similarly, the scattering from a microscale mesa can be modelled by the diffraction of two outward-facing wedge corners (Figure 3.1b). Incident rays are reflected by both the top and side surfaces (red and green arrows) and diffracted at the two wedge corners. Although the waves diffracted from a half-plane and a wedge are both cylindrical waves, the diffraction patterns are different and depend heavily on the wedge angle. This will be further discussed Section 3.2.

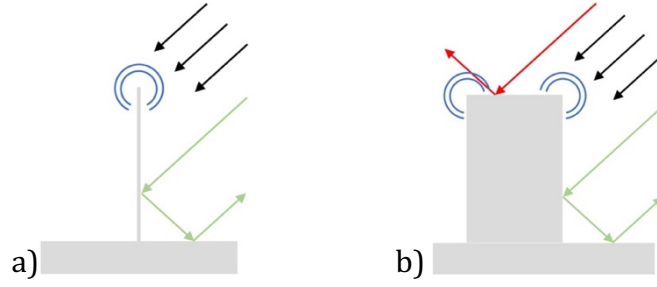


Figure 3.1: Diffraction from a) a half-plane and b) a mesa.

Since the amplitude of the diffracted field depends on the angle of illumination, Sommerfeld's half-plane problem is often solved in polar coordinates, and conversion from Cartesian coordinates is therefore needed. Figure 3.2 shows a Cartesian coordinate system. The half-plane is shown in red and is semi-infinite along the $-x$ -axis. The incident wave makes an angle α and the diffracted wave an angle θ to the $+z$ -axis. Cartesian coordinates can then be translated into polar expressions as $z = r \cos \alpha$ and $x = r \sin \alpha$.

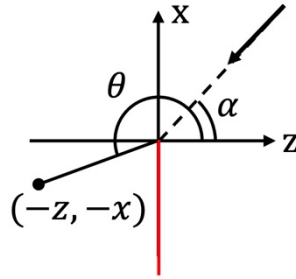


Figure 3.2: Sommerfeld's half-plane problem in Cartesian and polar coordinate systems.

Solution of the boundary matching problem then proceeds as follows. Recall the wave equation for a two-dimensional problem discussed in Section 2.2:

$$\frac{\partial^2 E_y}{\partial z^2} + \frac{\partial^2 E_y}{\partial x^2} + k^2 E_y = 0 \quad (3.1)$$

Here E_y is the total electric field for TE polarisation, conventionally written in terms of the incident E_i , reflected E_r and diffracted fields E_d . A similar wave equation exists for the magnetic field H_y for TM polarisation. For a perfect conductor, the boundary conditions for TE and TM polarization are:

$$E_y = 0 \quad (3.2)$$

$$\frac{\partial H_y}{\partial n} = 0 \quad (3.3)$$

Here $\partial/\partial n$ is the derivative along the normal to the boundary. Equations 3.2 and 3.3 are often described as Dirichlet and Neumann boundary conditions or as soft and hard

conditions. Here we assume TE-polarisation since the solution for TM-polarisation can be derived in the same fashion. A solution of Equation 3.1 for the incident plane wave can be written in Cartesian and polar coordinates as:

$$E_y = E_0 e^{+jk(z \cos \alpha + x \sin \alpha)} \quad E_y = E_0 e^{+jkr \cos(\theta - \alpha)} \quad (3.4)$$

For simplicity, the amplitude of the incident field E_0 is assumed as unity.

The total solution is then expressed as an angular spectrum of plane waves, written in polar coordinates as [7]:

$$E_y = \int f(\alpha) e^{+jkr \cos(\theta - \alpha)} d\alpha \quad (3.5)$$

Here the function $f(\alpha)$ is the angular spectrum of E_y . With a suitable integration path, $f(\alpha)$ can be found by Fourier transformation.

For example, the electric field can be formulated from the current density. The electric and magnetic field of a single TE-polarised plane wave can be written as:

$$\mathbf{E} = (0, 1, 0) e^{+jkr \cos(\theta - \alpha)} \quad \mathbf{H} = (\cos \alpha, 0, -\sin \alpha) e^{+jkr \cos(\theta - \alpha)} \quad (3.6)$$

Given the relationship between current density and the magnetic field ($H_z = -\frac{2\pi}{c} J_y$), the current density at point $(\xi, 0)$ must then be:

$$J_y(\xi) = +\frac{c}{2\pi} e^{+jk\xi \cos \alpha} \sin \alpha \quad (3.7)$$

The current density can also be rewritten as a Fourier integral. Making the substitution $\mu = \cos \alpha$, this integral can be written as:

$$J_y(\xi) = +\frac{c}{2\pi} \int_{-\infty}^{+\infty} P(\mu) e^{+jk\xi\mu} d\mu \quad (3.8)$$

The integrand here is the magnetic field H_z written in angular spectrum form, and the unknown function $P(\mu)$ is the angular spectrum. The electric field E_y can therefore also be written in terms of α or μ , as:

$$E_y = \int_C P(\cos \alpha) e^{+jkr \cos(\theta - \alpha)} d\alpha \quad E_y = \int_{-\infty}^{+\infty} \frac{P(\mu)}{\sqrt{1 - \mu^2}} e^{+jkz\mu} d\mu \quad (3.9)$$

The denominator $\sqrt{1 - \mu^2} = \sin \alpha$ is due to the $\sin \alpha$ term in H_z . The integration path C is in the complex α plane along which the real value of $\cos \alpha$ ranges from $+\infty$ to $-\infty$. This cancels out with the negative sign from differentiating $\cos \alpha$.

If a scattered field E_{ys} in TE-polarisation is represented as an angular spectrum of plane waves, dual integral equations can be formed using the boundary conditions, as:

$$E_{ys} = \int_{-\infty}^{+\infty} \frac{P(\mu)}{\sqrt{1-\mu^2}} e^{+jkz\mu} d\mu = -E_{yi} \quad (3.10)$$

$$0 = \int_{-\infty}^{+\infty} P(\mu) e^{+jkz\mu} d\mu \quad (3.11)$$

We now consider the TE-polarised plane wave incident on the perfectly conducting half-plane as $E_{yi} = e^{+jkr \cos(\theta-\alpha_0)}$, with the assumption of incidence at an angle $0 < \alpha_0 < \pi$ for simplicity. The subscript “0” denotes the incident field. The dual integral equations can be evaluated with standard contour integration techniques to integrate along a path in the complex plane without passing through any branch-points or singularities. Details of the necessary steps are omitted here. The function $P(\mu)$ is resolved and expressed as:

$$P(\mu) = -\frac{j}{2\pi} \frac{\sqrt{(1-\mu_0)}\sqrt{(1-\mu)}}{\mu + \mu_0} \quad (3.12)$$

Here $\mu_0 = \cos \alpha_0$. Similarly, the function $P(\cos \alpha)$ becomes:

$$P(\cos \alpha) = -\frac{j \sin \frac{1}{2} \alpha_0 \sin \frac{1}{2} \alpha}{\pi \cos \alpha + \cos \alpha_0} \quad (3.13)$$

Substitution back into Equation 3.10 then allows the scattered field to be obtained as:

$$E_y = e^{+jkr \cos(\theta-\alpha)} + \frac{1}{j\pi} \int_c \frac{\sin \frac{1}{2} \alpha_0 \sin \frac{1}{2} \alpha}{\cos \alpha + \cos \alpha_0} e^{-jkr \cos(\theta-\alpha)} d\alpha \quad (3.14)$$

Unfortunately, the complex integration representing the scattered field is too complicated for practical application, so simplification is needed.

When the observation point is further than one wavelength from the tip of the half-plane ($kr \gg 1$), Equation 3.14 can be evaluated by the method of steepest descent along the steepest descent $S(\theta)$ through the saddle-point at $\alpha = \theta$. First, the function $P(\cos \alpha)$ from Equation 3.13 is rearranged to obtain:

$$P(\cos \alpha) = -\frac{j}{4\pi} \left(\sec \frac{1}{2} (\alpha - \alpha_0) - \sec \frac{1}{2} (\alpha + \alpha_0) \right) \quad (3.15)$$

For simplicity, consider only the $(\alpha - \alpha_0)$ term in $P(\cos \alpha)$, as the second term can be obtained by changing the sign of α_0 . Let $P(\cos \alpha) = \sec \frac{1}{2} (\alpha - \alpha_0)$ in Equation 3.9 and we integrate along the steepest descent path $S(\theta)$ to obtain:

$$E_y = \int_{S(\theta)} \sec \frac{1}{2}(\alpha - \alpha_0) e^{-jkr \cos(\theta - \alpha)} d\alpha \quad (3.16)$$

We now define a new variable $\tau = \sqrt{2}e^{-\frac{j\pi}{4}} \sin \frac{1}{2}(\alpha - \theta)$. Transforming Equation 3.16 to integrate along $S(0)$, the result can be written in terms of τ as:

$$E_y = -2e^{-\frac{j\pi}{4}} e^{-jkr} \eta \int_{-\infty}^{\infty} \frac{e^{-kr\tau^2}}{\tau^2 + j\eta^2} d\tau \quad (3.17)$$

Here $\eta = \sqrt{2} \cos \frac{1}{2}(\theta - \alpha_0)$. Equation 3.17 can be further simplified in the form of the complex Fresnel integral $F(a) = \int_a^{\infty} e^{-j\tau} d\tau$, expressed as:

$$E_y = -2e^{-\frac{j\pi}{4}} e^{-jkr} \left(\pm 2\sqrt{\pi} e^{+jkr\eta^2} F(\pm\eta\sqrt{kr}) \right) \quad (3.18)$$

The upper (+) sign is for $\eta > 0$ and the lower (−) sign is for $\eta < 0$. The field expression when $P(\cos \alpha) = \sec \frac{1}{2}(\alpha + \alpha_0)$ equals Equation 3.18 with $\eta = \sqrt{2} \cos \frac{1}{2}(\theta + \alpha_0)$. Eventually, the scattered field term in Equation 3.14 can be rewritten as:

$$\begin{aligned} E_{ys} &= + \frac{1}{j\pi} \int_c \frac{\sin \frac{1}{2}\alpha_0 \sin \frac{1}{2}\alpha}{\cos \alpha + \cos \alpha_0} e^{-jkr \cos(\theta - \alpha)} d\alpha \\ &= -\frac{e^{\frac{j\pi}{4}}}{\sqrt{\pi}} \left\{ e^{jkr \cos(\theta - \alpha_0)} F \left[\sqrt{2kr} \cos \frac{1}{2}(\theta - \alpha_0) \right] \mp e^{jkr \cos(\theta + \alpha_0)} F \left[\pm \sqrt{2kr} \cos \frac{1}{2}(\theta + \alpha_0) \right] \right\} \end{aligned} \quad (3.19)$$

Here, the upper (+) sign is for $\theta + \alpha_0 < \pi$ and lower (−) is for $\theta + \alpha_0 > \pi$. The remaining simple pole $\alpha = \pi - \alpha_0$ in the integral in Equation 3.14 can be resolved and the contribution obtained from the residue theorem as:

$$-e^{+jkr \cos(\theta + \alpha_0)} \quad (3.20)$$

This is the reflected wave from geometrical optics (GO) and the discontinuity across $\theta = \pi - \alpha_0$ counterbalances that in the diffracted field in Equation 3.19. The complete solution (Equation 3.14) of the wave equation can then be written as:

$$E_y = \frac{E_0 e^{\frac{j\pi}{4}}}{\sqrt{\pi}} \left\{ e^{jkr \cos(\phi - \phi_i)} F \left[-\sqrt{2kr} \cos \left(\frac{\phi - \phi_i}{2} \right) \right] - e^{jkr \cos(\phi + \phi_i)} F \left[-\sqrt{2kr} \cos \left(\frac{\phi + \phi_i}{2} \right) \right] \right\} \quad (3.21)$$

$$H_y = \frac{E_0 e^{\frac{j\pi}{4}}}{\eta \sqrt{\pi}} \left\{ e^{jkr \cos(\phi - \phi_i)} F \left[-\sqrt{2kr} \cos \left(\frac{\phi - \phi_i}{2} \right) \right] + e^{jkr \cos(\phi + \phi_i)} F \left[-\sqrt{2kr} \cos \left(\frac{\phi + \phi_i}{2} \right) \right] \right\} \quad (3.22)$$

Equations 3.21 and 3.22 are the TE and TM solutions and are commonly known as Sommerfeld's solution [8]. Their properties can be explained as follows.

Figure 3.3a shows the Sommerfeld problem in polar coordinates once again. The elevation angle ϕ_i and the diffraction angle ϕ are now measured from the half-plane itself, while r is the distance from the tip of the plane to the observation point as before. Figure 3.3b shows the subspaces of the problem. The space is divided into three regions depending on the angle of incidence ϕ_i . Since the half-plane is fully reflective, the incident field does not travel in the shadow region (A). Likewise, incident field can only be reflected by the half-plane, so the reflected field only propagates in the reflection region (C). Since the cylindrical diffracted field radiates in all direction ($0 \leq \phi \leq 2\pi$), it exists in the entire space. As a result, the total electric field in each region is different.

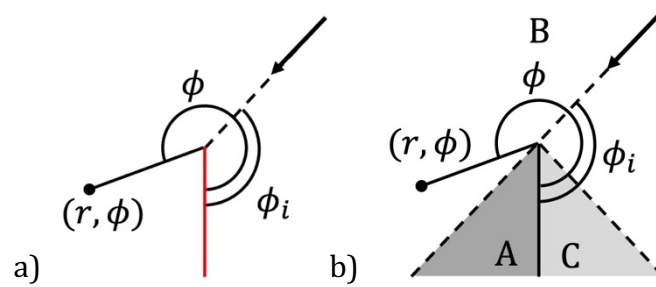


Figure 3.3: a) Sommerfeld's problem and b) the subspaces.

Sommerfeld's solution has since become crucial to most of the approaches to diffraction problems as it provides the only known closed form solution. However, its derivation was originally explained only briefly. Since then, many researchers have attempted to refine the solution and obtain simpler expressions for diffracted field [9-13]. We now briefly describe some of these attempts.

3.2 Geometrical theory of diffraction (GTD)

The geometrical theory of diffraction (GTD) was first proposed by Keller in 1962 [14, 15]. GTD describes the diffracted field by introducing additional diffracted rays to GO. Application of Fermat's principle shows that a ray from a point A to point B diffracted from an edge will take the shortest path. Figure 3.4 shows the Fermat's principle for edge diffraction on a half-plane. If a ray is incident from point A onto the edge of the half-plane at point C, the shortest path towards point B is the diffracted ray. The dashed lines are not the shortest path, so they are not considered as diffracted rays.

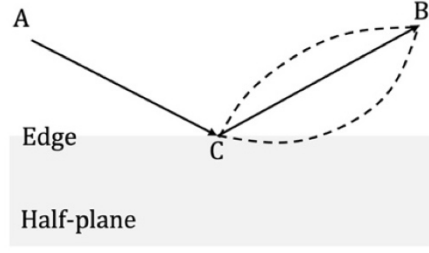


Figure 3.4: Illustration of Fermat's Principle for edge diffraction.

Figure 3.5 shows multiple diffracted rays generated by an incident ray that strikes the edge of the half-plane obliquely. This is known as Keller's cone. It implies that the wave diffracted from an edge is conical for oblique incidence. However, if the incident ray strikes the edge at normal incidence, the diffracted rays are also normal to the edge and radiate as cylindrical waves [16-19].

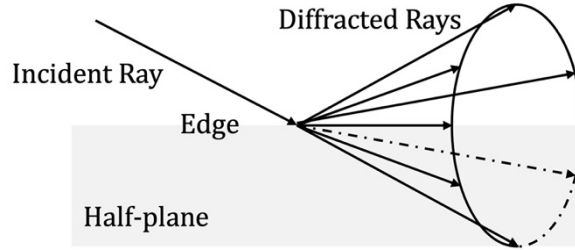


Figure 3.5: Keller's cone of diffracted rays.

Keller formulated a simple solution to the half-plane problem in the far field by retaining only the leading term in the asymptotic expansion of the Fresnel integral in Sommerfeld's solution. Now, the full asymptotic expansion for $a > 0$ is:

$$F(a) = -\frac{j}{2a}e^{-ja^2} + \frac{1}{4a^3}e^{-ja^2} - \frac{3}{4} \int_a^\infty \tau^{-4}e^{-j\tau^2} d\tau \quad (3.23)$$

To proceed further, we also need an identity property of the Fresnel integral:

$$F(a) + F(-a) = \sqrt{\pi}e^{-\frac{j\pi}{4}} \quad (3.24)$$

Substituting the first term in the Fresnel integral and the identity property into the Sommerfeld's solution (Equation 3.21) yields:

$$E = \frac{E_0 e^{\frac{j\pi}{4}}}{\sqrt{\pi}} \left\{ e^{jkr \cos(\phi - \phi_i)} \left[\sqrt{\pi} e^{-\frac{j\pi}{4}} + \frac{j}{2 \left[\sqrt{2kr} \cos\left(\frac{\phi - \phi_i}{2}\right) \right]} e^{-j \left[\sqrt{2kr} \cos\left(\frac{\phi - \phi_i}{2}\right) \right]^2} \right] \right. \\ \left. - e^{jkr \cos(\phi + \phi_i)} \left[\sqrt{\pi} e^{-\frac{j\pi}{4}} + \frac{j}{2 \left[\sqrt{2kr} \cos\left(\frac{\phi + \phi_i}{2}\right) \right]} e^{-j \left[\sqrt{2kr} \cos\left(\frac{\phi + \phi_i}{2}\right) \right]^2} \right] \right\} \quad (3.25)$$

Expanding the terms inside the curly bracket and rearranging the terms for the plane waves and edge waves gives,

$$E = \frac{E_0 e^{\frac{j\pi}{4}}}{\sqrt{\pi}} \left\{ e^{jkrcos(\phi-\phi_i)} \sqrt{\pi} e^{-\frac{j\pi}{4}} - e^{jkrcos(\phi+\phi_i)} \sqrt{\pi} e^{-\frac{j\pi}{4}} + \left[\frac{j}{\sqrt{8kr}} \sec\left(\frac{\phi-\phi_i}{2}\right) e^{jkrcos(\phi-\phi_i)} e^{-j\left[2kr\cos^2\left(\frac{\phi-\phi_i}{2}\right)\right]} - \frac{j}{\sqrt{8kr}} \sec\left(\frac{\phi+\phi_i}{2}\right) e^{jkrcos(\phi+\phi_i)} e^{-j\left[2kr\cos^2\left(\frac{\phi+\phi_i}{2}\right)\right]} \right] \right\} \quad (3.26)$$

Using the double angle formula, the exponential terms inside the square bracket can then be simplified to:

$$e^{jkrcos(\phi-\phi_i)} e^{-j\left[2kr\cos^2\left(\frac{\phi-\phi_i}{2}\right)\right]} = e^{jkrcos(\phi-\phi_i)} e^{-jkr[\cos(\phi-\phi_i)+1]} = e^{-jkr} \quad (3.27)$$

Keller's expression can then be obtained by simplifying Equation 3.26 to get:

$$E = E_0 e^{jkrcos(\phi-\phi_i)} - E_0 e^{jkrcos(\phi+\phi_i)} + \frac{E_0 e^{-\frac{j\pi}{4}} e^{-jkr}}{\sqrt{8\pi kr}} \left[-\sec\left(\frac{\phi-\phi_i}{2}\right) + \sec\left(\frac{\phi+\phi_i}{2}\right) \right] \quad (3.28)$$

This result has a simple interpretation, which has formed the basis of much high frequency theory. The first and second terms describe the incident and reflected plane waves respectively. The last term represents the diffracted wave, by assuming the diffracted field in cylindrical form as:

$$E_d = \frac{E_0 D e^{-jkr}}{\sqrt{r}} \quad (3.29)$$

Here D is the diffraction coefficient, an extremely important parameter. Comparison with Equation 3.28, allows the half-plane diffraction coefficient to be extracted as:

$$D = \frac{e^{-\frac{j\pi}{4}}}{\sqrt{8\pi k}} \left[-\sec\left(\frac{\phi-\phi_i}{2}\right) + \sec\left(\frac{\phi+\phi_i}{2}\right) \right] \quad (3.30)$$

For a wedge, the diffraction coefficient can be found using a similar (and equally lengthy) procedure, as:

$$D = \frac{e^{-\frac{j\pi}{4}} \sin \frac{\pi}{n}}{n\sqrt{2\pi k} \sin \beta} \left[\left(\cos \frac{\pi}{n} - \cos \frac{\phi-\phi_i}{n} \right)^{-1} \mp \left(\cos \frac{\pi}{n} - \cos \frac{\phi+\phi_i}{n} \right)^{-1} \right] \quad (3.31)$$

Here n is the coefficient of the exterior angle of the wedge, normally written as $(2-n)\pi$, and β is the azimuth angle of the illumination. For a half-plane ($n=2$) with perpendicular incidence ($\beta = \frac{\pi}{2}$), Equation 3.31 reduces to Equation 3.30. Figure 3.6

shows the diffraction pattern of a) a half-plane and b) a wedge for near field at $\theta = 50^\circ$, computed from Equations 3.30 and 3.31 respectively. The incident angle from normal is defined as θ and is commonly used for imaging. The diffraction pattern of a half-plane is symmetric and the value approaches infinity at the shadow boundaries due to the secant function. For wedge, the diffraction pattern is significantly different and will be discussed in the next chapter.

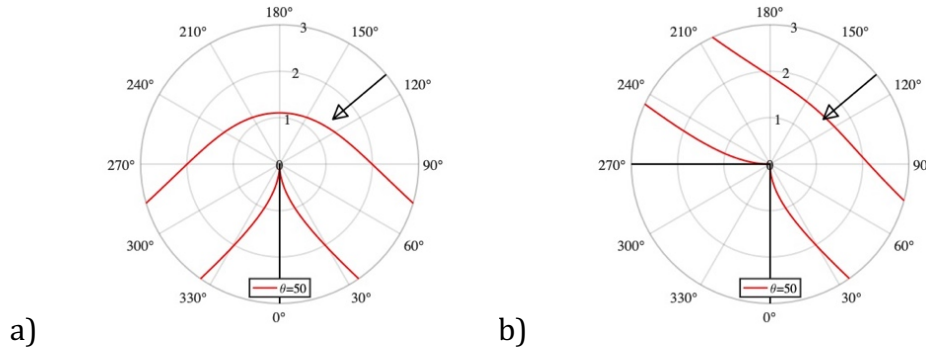


Figure 3.6: Diffraction pattern of a) a half-plane and b) a right-angled wedge.

In general, GTD calculates the total field accurately in the far field ($kr \gg 1$), since the higher order terms diminish rapidly. In this regime, GTD offers a close approximation to Sommerfeld's solution [20]. However, since the Fresnel integral $F(a)$ expands asymptotically with $\frac{1}{a^2}$ and $a = \sqrt{2kr} \cos\left(\frac{\phi + \phi_i}{2}\right)$, higher order terms are more significant when r is small, leading to inaccuracy in the near field. This leads to obvious discrepancies, especially at the boundaries between the regions in Figure 3.3b.

We illustrate this with a numerical example of a thin PEC mesa of width $0.1\mu m$ and height $5\mu m$ in a $10\mu m$ square simulation space. The plane wave illumination is at $\lambda = 0.5\mu m$ and incident angle of $\theta = 20^\circ$ from the y-axis. Note that the angle $\theta = \pi - \phi_i$ is used here, to simplify comparison with later simulations of image formation. Figure 3.7a and b show the total electric field from Sommerfeld's solution (Equation 3.21) and Keller's approximation (Equation 3.28) respectively, both computed numerically in MATLAB. Sommerfeld's solution is continuous across the shadow boundaries. However, Keller's approximation fails to match at the shadow boundary, and when the field strength increases along the reflection boundary. The region where the total field deviates from the exact solution is referred as the transition region. A magnitude limit of 5 is applied to

the secant function in Equation 3.28 to restrict the effects of the singularity in the transition region. This will be further discussed in the next section.

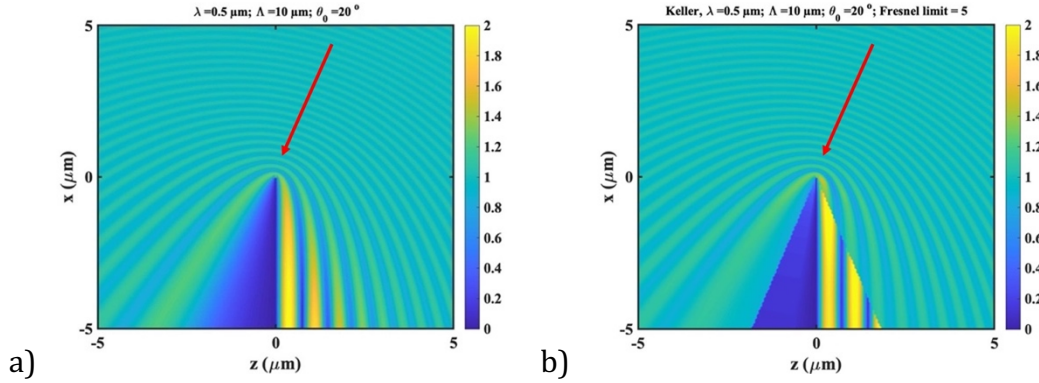


Figure 3.7: Total electric fields of a half-plane with incident angle of 20° , obtained from a) Sommerfeld's solution and b) Keller's approximation.

3.2.1 Fresnel integral

In general, the Fresnel integral is expressed as $F(a) = \int_0^a e^{-j\tau^2} d\tau$ [21]. The Fresnel integral can also be written in terms of its real and imaginary parts. The conversion between the exponential and trigonometric forms is:

$$F(a) = \int_0^a e^{-j\tau^2} d\tau = \int_0^a \cos(\tau^2) d\tau - j \int_0^a \sin(\tau^2) d\tau = C(a) - jS(a) \quad (3.32)$$

Figure 3.8a shows the Fresnel integral in trigonometric form. The red line represents the cosine integral $C(a)$ and blue line represents the sine integral $S(a)$. Both functions are decreasing towards $\pm\sqrt{\pi/8}$, when a tends to infinity [22]. Figure 3.7b is a parametric plot known as the Cornu spiral, showing the Fresnel integral on an Argand diagram.

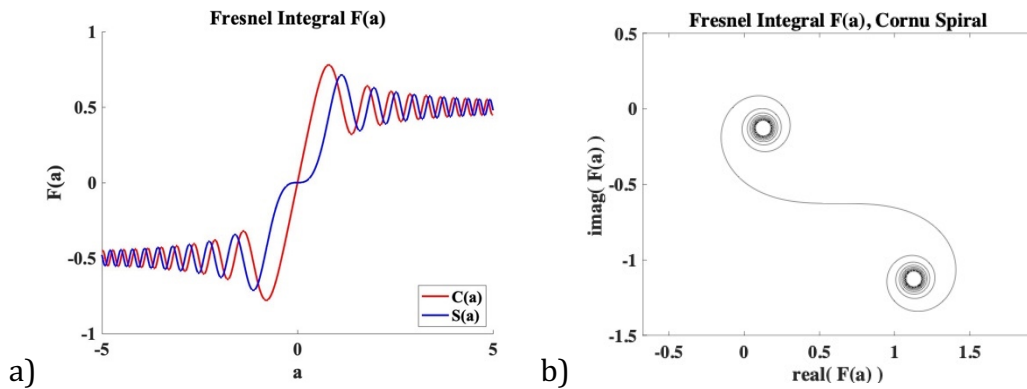


Figure 3.8: a) Variations of the Fresnel integrals in trigonometric form, and b) Cornu spiral.

In diffraction theory, particularly Sommerfeld's problem, the Fresnel integral is often defined alternatively as:

$$F(a) = \int_a^{+\infty} e^{-j\tau^2} d\tau \quad (3.33)$$

for $a > 0$. If we now introduce a new parameter b such that $b < 0$ and $b = -a$, the Fresnel integral can be written as the sum of two integrals, one giving the integral on the negative axis and the other on the positive axis:

$$F(b) = \int_b^0 e^{-j\tau^2} d\tau + \int_0^{+\infty} e^{-j\tau^2} d\tau \quad (3.34)$$

Due to the odd function characteristic of Fresnel integral as shown in Figure 3.7a, the "negative" integral can be written as:

$$\int_b^0 e^{-j\tau^2} d\tau = \int_{-\infty}^0 e^{-j\tau^2} d\tau - \int_{-\infty}^b e^{-j\tau^2} d\tau = \int_0^{+\infty} e^{-j\tau^2} d\tau - \int_a^{+\infty} e^{-j\tau^2} d\tau \quad (3.35)$$

The second term in Equation 3.34 is the asymptotic limit of the Fresnel integral:

$$F(0) = \int_0^{+\infty} e^{-j\tau^2} d\tau = \sqrt{\frac{\pi}{8}} (1 - j) = \frac{\sqrt{\pi}}{2} e^{-\frac{j\pi}{4}} \quad (3.36)$$

Therefore, when $a < 0$, the Fresnel integral can be expressed as:

$$F(-a) = \sqrt{\pi} e^{-\frac{j\pi}{4}} - F(a) \quad (3.37)$$

Because the Fresnel integral is a transcendental function and cannot be evaluated in a closed form, it can only be resolved by asymptotic expansion or substitution. Asymptotic expansion is carried out by multiple integration by parts with respect to $\frac{1}{a}$. Keller's GTD is the simplest example of the asymptotic expansion. However, many researchers have proposed simpler substitutions [23], particularly James [24] and Abuelma'Atti [25].

Asymptotic expansion is the most straightforward approach, and the first four terms of the expansion are:

$$F(a) = \int_a^{\infty} e^{-j\tau^2} d\tau = e^{-ja^2} \left(-\frac{j}{2a} + \frac{1}{4a^3} + \frac{3j}{8a^5} - \frac{15}{16a^7} + \dots \right) \quad (3.38)$$

If we now consider only the first two terms in the expansion and substitute them into Sommerfeld's solution (Equation 3.21) for the diffracted field, we obtain:

$$E = E_{const} \left[\frac{j}{2\sqrt{2kr} \cos\left(\frac{\phi - \phi_i}{2}\right)} - \frac{1}{4\left(\sqrt{2kr} \cos\left(\frac{\phi - \phi_i}{2}\right)\right)^3} - \frac{j}{2\sqrt{2kr} \cos\left(\frac{\phi + \phi_i}{2}\right)} + \frac{1}{4\left(\sqrt{2kr} \cos\left(\frac{\phi + \phi_i}{2}\right)\right)^3} \right] \quad (3.39)$$

Here we focus on the Fresnel integral terms, grouping the other constants as a constant E_{const} . Further factorising Equation 3.39 to obtain a similar expression to Keller's expression (Equation 3.28) we get:

$$E = \frac{jE_{const}}{2\sqrt{2kr}} \left[\sec\left(\frac{\phi - \phi_i}{2}\right) - \sec\left(\frac{\phi + \phi_i}{2}\right) \right] \left\{ 1 + \frac{j}{4kr} \left[\sec^2\left(\frac{\phi - \phi_i}{2}\right) + \sec\left(\frac{\phi - \phi_i}{2}\right) \sec\left(\frac{\phi + \phi_i}{2}\right) + \sec^2\left(\frac{\phi + \phi_i}{2}\right) \right] \right\} \quad (3.40)$$

Retaining an extra term in the Fresnel integral expansion adds a complicated trigonometric term to Keller's expression. Correction of the Keller's expression improves the phase mismatch at the boundary region in Figure 3.7b. However, the correction terms also cause a drastic increase of field strength around the feature tip and at the boundaries. Setting a limit to the secant function can "force" an improvement in the matching of the magnitude of the electric field on the illuminated side of the feature, but the boundary matching in the shadow region is still inaccurate. Figure 3.9 shows the total electric field computed for $\theta = 20^\circ$ with the asymptotic expansion of the first four terms of the Fresnel integral imposing limits to the secant function of a) 20 and b) 10. There is still noticeable difference at the shadow boundaries in terms of magnitude and phase.

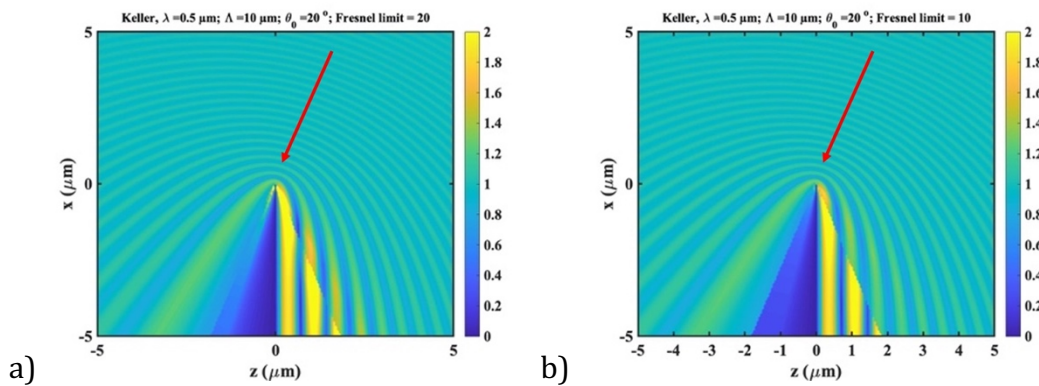


Figure 3.9: Total electric field at $\theta = 20^\circ$, computed from the asymptotic expansion of the Fresnel integral with secant limits of a) 20 and b) 10.

Theoretically, the more terms are included in the Fresnel integral expansion, the closer the agreement with Sommerfeld's solution. However, the correction terms become more

complicated, and the simplicity of Keller's approximation is lost. An alternative is to replace the Fresnel integral with a closed-form expression. Particularly, James' approximation based on Logan's introduction of a K expression [26] states that:

$$F(x) = \sqrt{\pi}K(x)e^{-j(x^2 + \frac{\pi}{4})} \quad (3.41)$$

James proposed an expression for $K(x)$ in the form:

$$K(x) = \frac{\exp\left[-j\left(\arctan(x^2 + 1.5x + 1) - \frac{\pi}{4}\right)\right]}{2\sqrt{\pi x^2 + x + 1}} \quad (3.42)$$

Here the arctan function is approximated as:

$$\arctan(x^2 + 1.5x + 1) = \frac{\pi}{2} - \frac{(x^2 + 1.5x + 1)}{0.28 + (x^2 + 1.5x + 1)^2} \quad (3.43)$$

Since the Cornu spiral indicates the magnitude change of the real and imaginary component of the Fresnel integral, it can be used to analyse the accuracy of the approximation. The error in amplitude and phase components of $K(x)$ from the exact value range from 2% to 8% for James' approximation. Abuelma'Atti then introduced a further correction term $0.022x \exp(-0.29x)$ to reduce the error to less than 3.4%.

These approximations are compared in Figure 3.10. Figure 3.10a shows the "perfect" Cornu spiral, representing the Sommerfeld's solution. Figures 3.10b, c and d show the spirals from Keller's, James' and Abuelma'Atti's approximations respectively. Figure 3.10b stands out by its lack of convergence for small a , implying that Keller's solution does not provide a good approximation in the near field. However, when a increases (for far field), Keller's Cornu spiral does approach the limit $\pm\sqrt{\pi/8}$. Keller's approximation is therefore accurate in the far field. However, the traces in Figure 3.10c and d follow the "ideal" Cornu spiral in Figure 3.10a closely, showing that these approximations will be generally accurate.

Figure 3.11 compares the total electric field obtained using the approaches discussed in Figure 3.10. Figures 3.11a-d show the electric field plots of Sommerfeld's solution and Keller's, James' and Abuelma'Atti's approximations for $\theta = 20^\circ$ (indicated as red arrows). As expected, the James' and Abuelma'Atti's approximations estimate the field accurately and the plots are very similar to Sommerfeld's solution.

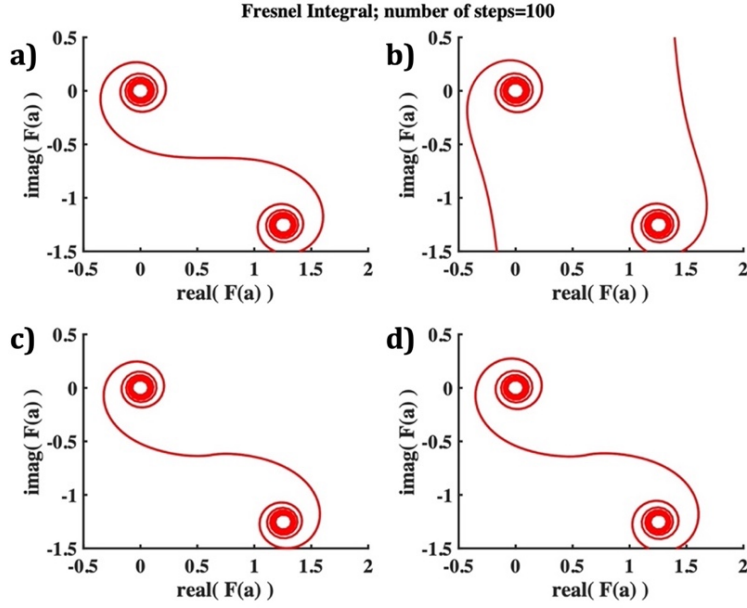


Figure 3.10: Cornu spiral for Fresnel integrals by a) Sommerfeld's, b) Keller's, c) James' and d) Abuelma'Atti's expressions.

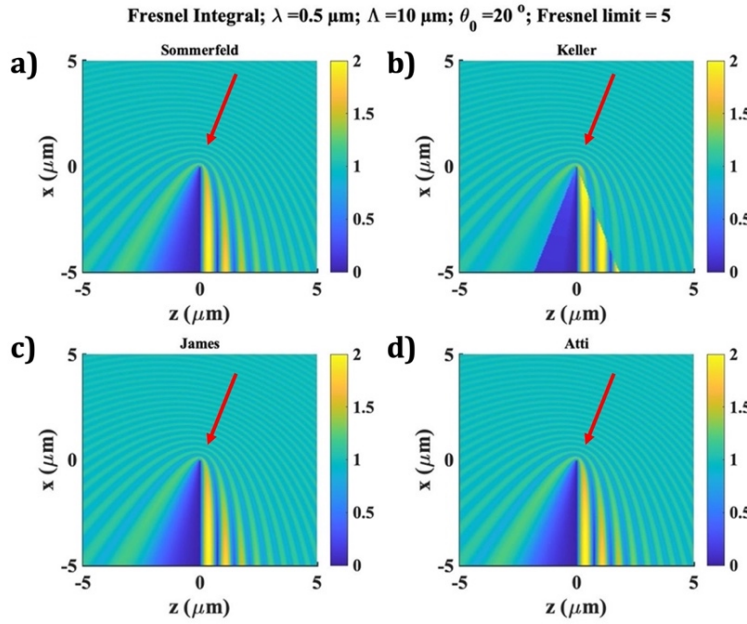


Figure 3.11: Total electric field computed using a) Sommerfeld's, b) Keller's, c) James' and d) Abuelma'Atti's solutions for $\theta = 20^\circ$.

3.2.2 Uniform theory of diffraction (UTD)

To reduce the singularity distortion of the electric field in GTD, two important uniform theories were created: the uniform theory of diffraction (UTD) and uniform asymptotic theory (UAT) [27-29]. The difference between the two theories is that UTD modifies the

diffracted field in GTD for the singularity in the transition region while UAT modifies the GO field in GTD to incorporate the singularity.

UTD scales the diffracted field according to the position of the observation point from the origin of the diffracted field. The diffracted field is redefined as E_D (with capital letter, different from the E_d in GTD) [30]. The full expression for E_D is:

$$E_D = \frac{F(|\xi_i|)}{\hat{F}(|\xi_i|)} E_{di} + \frac{F(|\xi_r|)}{\hat{F}(|\xi_r|)} E_{dr} \quad (3.44)$$

Here $\xi_{i,r}$ represents the incidence (reflection) detour parameter, which characterises how close the observation point is to the shadow boundaries, and is defined as $\xi_{i,r} = \epsilon_{i,r} \sqrt{k|S_d - S_{i,r}|}$. S_i is the direct distance from the source to the observation point and S_r (S_d) is the distance from the source to the observation point via the reflection (diffraction) point. $F(x)$ is a transition function from the Fresnel integral and $\hat{F}(x) = \frac{e^{-j(x^2 + \frac{\pi}{4})}}{2\sqrt{\pi x}}$. Using this approach, the equivalent diffraction coefficient to Equation 3.31 is:

$$D_{s,h} = \frac{-e^{-\frac{j\pi}{4}}}{2n\sqrt{2\pi k} \sin \beta} \times \left\{ \cot\left(\frac{\pi + (\phi - \phi_i)}{2n}\right) F(kra^+(\phi - \phi_i)) + \cot\left(\frac{\pi - (\phi - \phi_i)}{2n}\right) F(kra^-(\phi - \phi_i)) \right. \\ \left. \mp \left[\cot\left(\frac{\pi + (\phi + \phi_i)}{2n}\right) F(kra^+(\phi + \phi_i)) + \cot\left(\frac{\pi + (\phi + \phi_i)}{2n}\right) F(kra^-(\phi + \phi_i)) \right] \right\} \quad (3.45)$$

Here the sign $\mp$ depends on the polarisation of the incidence fields. The distinction to GTD arises from the transition function $F(kra)$, which is obtained from a uniform asymptotic evaluation of the wedge Green's function [30]. When the observation point is further away from the shadow boundaries ($kra \gg 0$), $F(kra) \rightarrow 1$ and the UTD coefficient reduces to Keller's result (Equation 3.32). This solves the singularity problem, but the caustic problem in GTD still remains [31].

Alternatively, the approach taken in UAT is to modify the GO field E_{go} to create a similar singularity that cancels out the shadow boundary singularity in GTD [32, 33]. The modified GO field E_{GO} is:

$$E_{GO} = [F(\xi_i) - \hat{F}(\xi_i)] E_i + [F(\xi_r) - \hat{F}(\xi_r)] E_r \quad (3.46)$$

The function $F(x)$ and $\hat{F}(x)$ are defined as in UTD but the sign of $\xi_{i,r}$ is important. The formulations of the two theories are different but numerically the results are similar. In practice, UTD has been the more popular method because of its simplicity. Due to the improvement in the transition region, UTD has become popular in more complex

diffraction problems [34, 35]. For example, multiple diffraction problems are often solved with UTD, as will be discussed later.

3.3 Physical theory of diffraction (PTD)

Despite the wide application of GTD and its improvements, UTD and UAT [36], there are still canonical problems that cannot be solved [37]. Physical optics (PO) techniques were therefore developed to find approximate diffraction coefficients. PO uses approximations of the induced surface current density on a perfectly conducting surface in proportion to the tangential incident magnetic field. However, it fails to evaluate the edge diffracted fields. An improved technique called the Physical Theory of Diffraction (PTD), modified from the Kirchhoff approximation (KA), was therefore developed [38-40]. The derivation of PTD is omitted here but the equations will be provided and compared with the solution from GTD. Figure 3.12 shows the geometry of PTD in a polar coordinate system and the half-plane (in red) is placed horizontally at $\phi = 0$. The arrow indicates the incident field from the source (r_0, ϕ_i) on to the surface $\phi = 0$, and the dot shows the observer at (r, ϕ) . The wedge angle is $2\pi - \alpha$ and the representation in the PTD formula is $n \left(n = \frac{\alpha}{\pi} \right)$.

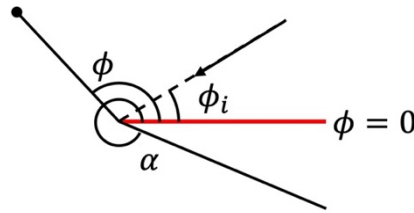


Figure 3.12: Geometry of PTD.

PTD separates the surface scattering sources into uniform and nonuniform components, designated using superscripts “0” and “1”. The uniform component j^0 is determined by GO, due to the current induced by the incident plane wave uniformly distributed on the surface. The nonuniform source j^1 is caused by smooth or sharp bends of the surface and is also responsible for fringe waves [38]. For TE-polarisation, the uniform current sources and the electric field generated by the uniform source can be expressed as:

$$j^0 = -E_0 2ik \sin \phi_i e^{-jkz \cos \phi_i} \quad E_d^0 = -\frac{1}{4\pi} \int_S j^0 \frac{e^{-jkr}}{r} ds \quad (3.47)$$

Equation 3.48 gives the approximate value of the scattered field from the PO approach [41]. The fringe wave E_{PTD} caused by the nonuniform source j^1 can be expressed as:

$$E_{PTD} = E_d - E_d^0 \quad (3.48)$$

Here the E_d is the exact solution of diffracted field and is expressed as a form of cylindrical wave:

$$E_d = E_0 f \frac{e^{-j(jr + \frac{\pi}{4})}}{\sqrt{2\pi kr}} \quad (3.49)$$

This is equivalent to the diffracted field in GTD, where f is the trigonometric part of the diffraction coefficient in GTD (Equation 3.31). If E_d^0 is expressed in the same form as Equation 3.49, the trigonometric term f_0 for the uniform diffracted field of a wedge with angle $(2\pi - \alpha)$ becomes:

$$f^0 = \frac{\sin \phi_i}{\cos \phi + \cos \phi_i} \quad f^0 = \frac{\sin \phi_i}{\cos \phi + \cos \phi_i} + \frac{\sin(\alpha - \phi_i)}{\cos(\alpha - \phi) + \cos(\alpha - \phi_i)} \quad (3.50)$$

The first part of Equation 3.50 is for $0 < \phi_i < \alpha - \pi$ and the second part is for $\alpha - \pi < \phi_i < \pi$, with an additional term due to the illumination of second surface of the wedge. Substituting Equation 3.50 into Equation 3.48, the integrand of the fringe wave E_{PTD} can be obtained by expanding the solution as a Taylor series and dropping higher order terms:

$$E_{PTD} = E_0 f^1 \frac{e^{-j(jr + \frac{\pi}{4})}}{\sqrt{2\pi kr}} + O\left[(kr)^{-\frac{3}{2}}\right] \quad (3.51)$$

The trigonometric term f^1 is the difference between f from the exact solution and f^0 in the uniform diffracted field. For near field diffraction, the second term in Equation 3.51 is dropped and the fringe wave becomes a cylindrical wave.

Figure 3.13 shows the fringe diffracted fields arising from the nonuniform source of a half-plane and a right-angled wedge at incident angle $\theta = 130^\circ$ for TE polarisation. The incident angle θ is again defined as $\theta = \pi - \phi_i$ for convenience, and it is measured from the horizontal plane at $\phi = 180^\circ$. The diffraction patterns of a half-plane and a wedge are different, as are the magnitudes. When the incident wave illuminates both surfaces on the wedge, the fringe effect is increased for TE-polarisation. For single side illumination, the fringe effect is much smaller, and the magnitude is halved or even smaller, depending on the incident angle and the wedge angle [42].

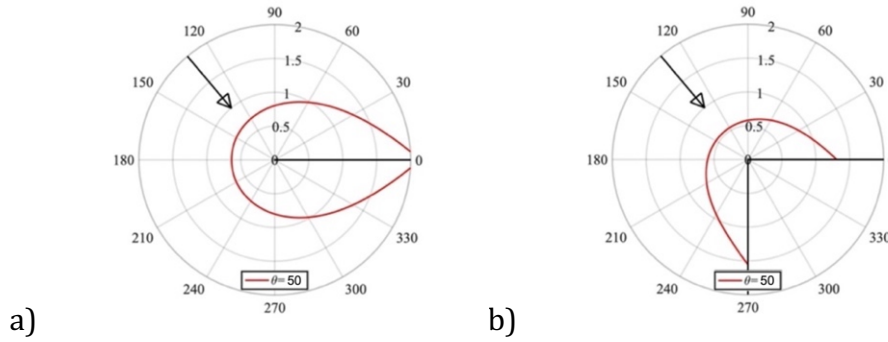


Figure 3.13: Fringe wave pattern of a) a half-plane and b) a right-angled wedge.

Despite the improvement offered by PTD, it only considers Fermat's principle for the PO current and is not valid at the edge due to the number of rays emerging from the tip of the wedge [43]. Further improvements have subsequently been made. One is the application of the equivalent edge current (EEC), which eliminates the caustic singularities in the original ray tracing PTD [44, 45]. Another is an extension for observation angles that are not on Keller's cone using incremental length diffraction coefficients (ILDC) [46-48]. A more complete modification is a derivation of the EEC by asymptotically reducing the surface-to-edge integral, written in terms of diffraction coefficients [49, 50]. The modified theory of physical optics (MTPO) proposed a new unit vector to dissect the angle between the incident and reflected rays into two equal parts. The edge of the scatterer is therefore modified, and the singularity in EEC for PO and GTD is removed [51]. However, these improvements will not be discussed in detail here. To summarize, PTD matches the results calculated from GTD and method of moments (MoM) [52] and is used as a numerically calculated reference solution [53-55].

3.4 Incremental theory of diffraction (ITD)

As mentioned above, incremental methods such as ILDC [46, 47, 56], EEC [48, 49] and elementary edge waves (EEW) [57, 58] that describe elementary contributions were used to eliminate caustic singularities and extend observation angles in PTD. A related method, the incremental theory of diffraction (ITD) was also developed to provide a way of computing diffraction from curved edges [59-61]. It introduces a localisation process based on constructing a suitable spatial integral representation of the 3D canonical Green's function through a 2D Green's function factorisation. It interprets the exact solution for the scattering in the near zone from a canonical cylindrical structure when it is illuminated by an elementary source.

Consider an object with a uniform cross-section along a curved line l . The diffracted $E_d(P)$ field at the observation point P is diffracted from the point Q_l when the object is illuminated by a point source P_0 . Figure 3.14 shows the geometry of a wedge with exterior angle α (or $n\pi$) in a spherical (r, β, ϕ) coordinate system [62].

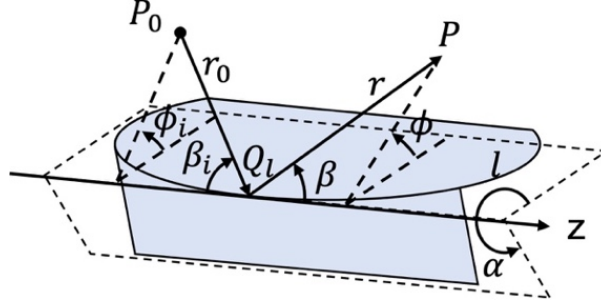


Figure 3.14: Geometry at locally tangent wedge.

The incremental field contribution $dE(Q_l)$ at P , which is localised at each point Q_l along l , is integrated along l such that the diffracted field $E_d(P)$ is expressed as:

$$E_d(P) = \int_l dE(Q_l) dl \quad (3.52)$$

Since it is assumed that scattering is a local phenomenon, the local segment of the object is approximated as a cylindrical canonical configuration with the z -axis tangent to the line l at Q_l . Hence, the exact solution E_c for the local problem is an infinite, spatial superposition of the elementary contributions $dE_c(z)$ along the z -axis:

$$E_c(P) = \int_{-\infty}^{+\infty} dE_c(z) dz \quad (3.53)$$

Here $dE(Q_l) = dE_c(z)$ at the tangential point $z = 0$. An exact spectral integral representation of the Green's function of an object with a uniform cylindrical canonical configuration along the z -axis can then be obtained as:

$$E_c(P) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G_{2d}(\rho, \rho'; k''_\rho) e^{-jk''_z(z-z')} dk''_z \quad (3.54)$$

Here $k''_\rho = \sqrt{k^2 - (k''_z)^2}$ and $G_{2d}(\rho, \rho'; k''_\rho)$ is the two-dimensional Green's function of the same object illuminated by a z -directed line source. In general, $G_{2d}(\rho, \rho'; k''_\rho)$ can be represented as the result of the application of a linear operator L to the product of two spectrum functions, each depending on the location of either the source $k''_\rho \rho'$ or the observation point $k''_\rho \rho$. Therefore, Equation 3.54 can be represented by the linear operator L and the spectrum functions U :

$$E_c(P) = L \left[\frac{1}{2\pi} \int_{-\infty}^{+\infty} U(k''_\rho, \rho) U(k''_\rho, \rho') e^{-jk''_z(z-z')} dk''_z \right] \quad (3.55)$$

From Fourier transform analysis, the integral can also be considered as a spatial convolution product of two functions, expressed as:

$$u(\zeta^{(')}, \rho^{(')}) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} U(k^{(')}_\rho, \rho^{(')}) e^{jk^{(')}_z \zeta^{(')}} dk^{(')}_z \quad (3.56)$$

The bracketed power (') indicates the quantities for unprimed $\zeta = z'' - z$ and primed $\zeta = z' - z''$, and the function denotes the inverse FT with respect to $k^{(')}_z$. Therefore, the incremental field contribution at Q_l at $z'' = 0$ is:

$$dE(Q_l) = L[u(-z, \rho) \cdot u(z', \rho')] \quad (3.57)$$

The above Fourier transform analysis is described as the ITD FT-convolution process and provides a general procedure for obtaining incremental field contributions from the Green's function of infinite, uniform cylindrical local canonical problems. For a wedge structure (in Figure 3.13), the dyadic and closed form expressions dE_d for the ITD incremental field contribution can be expressed as [63]:

$$dE_d(P) = \begin{bmatrix} dE_{d,\beta}(P, Q_l) \\ dE_{d,\phi}(P, Q_l) \end{bmatrix} = \begin{bmatrix} D_e & 0 \\ 0 & D_h \end{bmatrix} \cdot \begin{bmatrix} E_{i,\beta}(Q_l) \\ E_{i,\phi}(Q_l) \end{bmatrix} \frac{e^{-jkr}}{2\pi r} \quad (3.58)$$

Here the $E_i(Q_l)$ is the incident field at Q_l and $D_{e,h}$ is the incremental diffraction coefficient. The total diffracted field observed at P is the superposition of the incremental contributions distributed along the contour L , expressed as a line integral as:

$$E_d(P) = \int_L dE_d(P) dl = \int_L D(l) E_0(l) \frac{e^{-jkr(l)}}{r(l)} dl \quad (3.59)$$

The total field at point P is simply the sum of $E_d(P)$ and the GO field E_{go} . ITD is simple to apply to any diffraction problem and complies with the UTD solution. Also, since only Equation 3.59 must be solved, computation of the diffracted field is much faster than other methods. However, a limitation of ITD is that the caustic problem remains [64]. Nonetheless, ITD is still widely applicable to many diffraction problems [65-68].

3.5 Multiple diffraction

In this chapter, diffraction from a single isolated half-plane or wedge corner has so far been discussed. However, MEMS and NEMS devices are usually composed of multiple closely placed mesas. A mesa can be considered as a pair of closely spaced knife edges or wedges [69-72]. A secondary diffracted field is then generated when a primary diffracted field impinges upon a nearby corner. Higher order diffraction is theoretically possible but

is usually disregarded due to its significantly smaller magnitude. GTD or UTD can usually be applied in such cases. However, the solution can be inaccurate when the edge separation is small compared to the wavelength, because of the rapid transition across the shadow boundaries, from one ray optical form in the lit region to another ray optical form in the shadow region [73]. One solution is the iterative surface current density replacement technique [74].

Figure 3.15 shows the diffraction from a two-dimensional double mesa configuration with a separation m between the mesas. The black, red, green arrows represent the incident, reflected and retro-reflected rays respectively. The blue cylindrical waves are the diffracted waves from each mesa corner. The orange arrows between the mesas are the doubly diffracted field (E_{dd}), whose amplitude is inversely proportional to the square root of the separation distance ($E_{dd} \propto \frac{1}{\sqrt{m}}$).

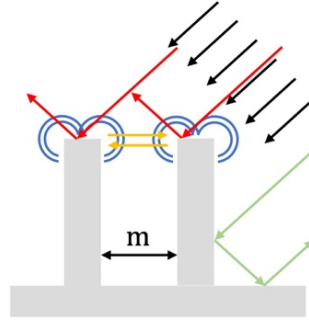


Figure 3.15: Diffraction from two closely spaced mesas.

Considering the wedge corners of the inner side of the mesas, the secondary diffracted field E_{dd} is defined as:

$$E_{dd} = E_0 D_1 D_2 A_{dd} e^{-jkr} \quad (3.60)$$

Here the spreading factor $A_{dd} = \frac{1}{\sqrt{rm}}$, implying that the secondary diffracted field is also a cylindrical wave. $D_{1,2}$ describes the diffraction coefficient from the incident field and the diffracted field on the inner corner of the first feature [75-77]. For a three-dimensional scattering problem, the two mesas are not always in parallel. This is taken into account by an extra factor of $\frac{1}{\sin \beta_{1,2}}$ in the diffraction coefficients D_1 and D_2 [78-80]. For multiple diffraction, Equation 3.60 can be extended to multiply the additional diffraction coefficients with the corresponding spreading factor [81, 82]. This simple method of

multiple diffraction is used later for the development of the novel analytic model of more general MEMS and NEMS imaging.

3.6 Summary

The famous Sommerfeld half-plane diffraction problem has been introduced, together with the derivation of the solution. Several approximations to Sommerfeld's solution have been demonstrated, including the ray optics and the surface current integration approaches. Despite the occasional inaccuracy of GTD in the transition and caustic regions, it estimates the diffracted field accurately for most scenarios, and especially at the large distances important in the MEMS and NEMS imaging application. In the next chapter, GTD will be compared to full-wave simulation carried out using the FEM for scattering from half-plane and wedge models. In particular, the accuracy of analytical predictions obtained using GTD for the scattering from a NEMS device will be validated using FEM results.

References

Historical

1. T. Young, "The Bakerian lecture: on the theory of light and colours", *Phil. Trans. Roy. Soc. Lond.*, vol. 92, pp. 12-48, 1802.
2. A. Fresnel, "Memoir on the diffraction of light", *Mém. Acad. Roy. Sci. Inst. France*, vol. 5, pp. 339-475, 1818.
3. G. R. Kirchhoff, "Zur theorie der lichtstrahlen", *Ann. Phys.*, vol. 18, pp. 663-695, 1882.
4. H. von Helmholtz, "Theorie der luftschwingungen in rohren mit offenen enden", *J. fur die Reine und Angew. Math.*, vol. 57, pp. 1-72, 1859.

Sommerfeld half-plane problem

5. A. J. W. Sommerfeld, "Mathematische theorie der diffraction", *Math. Ann.*, vol. 47, pp. 317-374, 1896.
6. A. J. W. Sommerfeld, "Optics", Academic Press, New York, 1954.
7. E. T. Whittaker, G. N. Watson, "A course of modern analysis", Cambridge University Press, Cambridge, 1927.
8. M. Born, E. Wolf, "Principles of optics: electromagnetic theory of propagation, interference and diffraction of light", Sixth Ed., Pergamon Press, Oxford, 1980.
9. G. D. Durgin, "Practical geometrical behavior of knife-edge diffraction", *Proc. IEEE Ant. Propag. Soc. Int. Symp.*, San Diego, CA, July 5-11, pp. 1-4, 2008.
10. E. V. Meister, "Generalized Sommerfeld half-plane problems", *Equadiff 5, Proc. Fifth Czechoslovak Conf. on Diff. Eq. and Their Apps.*, Bratislava, Slovakia, Aug. 24-28, pp. 264-267, 1982.
11. A. E. Heins, A. Arbor, "The Sommerfeld half-plane problem revisited I: the solution of a pair of coupled Wiener-Hopf integral equations", *Math. Meth. Appl. Sci.*, vol. 4, pp. 74-90, 1982.
12. F. Gori, "Diffraction from a half-plane. A new derivation of the Sommerfeld solution", *Opt. Comm.*, vol. 48, pp. 67-70, 1983.
13. R. W. Ziolkowski, "Exact solution of the Sommerfeld half-plane problem: a path integral approach without discretization", *J. Elect. Waves Appl.*, vol. 1, pp. 377-402, 1987.

Geometrical theory of diffraction

14. J. B. Keller, "Diffraction by an aperture", *J. Appl. Phys.*, vol. 28, pp. 426-444, 1957.

15. J. B. Keller, "Geometrical theory of diffraction" J. Opt. Soc. Am., vol. 52, pp. 116-130, 1962.
16. T. B. A. Senior, P. L. E. Uslenghi, "Experimental detection of edge-diffraction cone", Proc. IEEE, vol. 60, pp. 1448, 1972.
17. R. K. Luneburg, "Propagation of electromagnetic plane waves in plane parallel layers", CIMS, New York University, New York, 1947.
18. M. Kline, "An asymptotic solution of Maxwell's equations", Comm. Pure Appl. Math., vol. 4, pp. 225-263, 1951.
19. T. B. A. Senior, P. L. E. Uslenghi "Comparison between Keller's and Ufimtsev's theories for the strip", IEEE Trans. Ant. Propag., vol. AP-19, pp. 557-558, 1971.
20. P.L. Christiansen, N.Chr. Albertsen, O. Breinbjerg, "50 years with J. B. Keller's geometrical theory of diffraction in Denmark – revisiting the theory: impedance half-plane diffraction coefficients" IEEE Ant. Propag. Mag., vol. 55, pp. 32-40, 2013.

Fresnel Integral

21. M. Sandoval-Hernandez, H. Vazquez-Leal, *et. al.*, "Approximation of Fresnel integrals with applications to diffraction problems", Math. Prob. in Eng., vol. 2018, pp. 1-13, 2018.
22. S. Anokhov, "Physical approach to analytic simulation of Fresnel integrals", J. Opt. Soc. Am. A., vol. 24, pp. 197-203, 2007.
23. Y. Z. Umul, "Equivalent functions for the Fresnel integral", Opt. Express, vol. 13, pp. 8469-8482, 2005.
24. G. James, "An approximation to the Fresnel integral", Proc. IEEE, vol. 67, pp. 677-678, 1979.
25. M. Abuelma'Atti, "An improved approximation to the Fresnel integral", IEEE Trans. Ant. Propag., vol. 37, pp. 946-947, 1989.
26. L. A. Logan, "General research in diffraction theory vol. 1", Publ. LMSD-288087, Missiles and Space Div., Lockheed Aircraft Corp., Sunnyvale, CA, 1959.

Uniform theory of diffraction and Uniform asymptotic theory

27. R. G. Kouyoumjian, P. H. Pathak, "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface" Proc. IEEE, vol. 62, pp. 1448-1461, 1974.
28. R. M. Lewis, J. Boersma, "Uniform asymptotic theory of edge diffraction" J. Math. Phys., vol. 10, pp. 2291-2305, 1969.
29. S. W. Lee, G. Deschamps, "A uniform asymptotic theory of electromagnetic

- diffraction by a curved wedge", IEEE Trans. Ant. Propag., vol. AP-24, pp. 25-34, 1976.
30. P. H. Pathak, "Uniform geometrical theory of diffraction" Proc. SPIE, vol. 0358, pp. 47-61, 1982.
 31. R. Paknys, "Applied frequency-domain electromagnetics", IEEE Press, New York, 2016.
 32. Y. Rahmat-Samii, R. Mittra, "Spectral analysis of high frequency diffraction of an arbitrary incident field by a half-plane comparison with four asymptotic techniques", Radio Sci., vol. 13, pp. 31-48, 1978.
 33. S. W. Lee, "Uniform asymptotic theory of electromagnetic edge diffraction: A review", Electromag. Scat., New York: Academic Press, New York, pp. 67-119, 1978.
 34. G. Carluccio, M. Albani, P. H. Pathak, "Uniform asymptotic evaluation of surface integrals with polygonal integration domains in terms of UTD transition functions", IEEE Trans. Ant. Propag., vol. 58, pp. 1155-1163, 2010.
 35. F. Puggelli, G. Carluccio, M. Albani, "An efficient algorithm for the computation of the UTD T transition function", IEEE Trans. Ant. Propag., vol. 60, pp. 2380-2387, 2012.
- Physical theory of diffraction
36. R. C. Hansen, "Geometric theory of diffraction", IEEE Press, New York, 1981.
 37. N. D. Taket, R. E. Burge, "A physical optics version of geometrical theory of diffraction", IEEE Trans. Ant. Propag., vol. 39, pp. 719-731, 1991.
 38. P. Y. Ufimtsev, "Approximate computation of the diffraction of plane electromagnetic waves at certain metal bodies" Sov. Phys. Tech. Phys., vol. 27, pp. 1708-1718, 1957.
 39. P. Y. Ufimtsev, "Secondary diffraction of electromagnetic waves by a strip" Sov. Phys. Tech. Phys., vol. 28, pp. 535-548, 1958.
 40. P. Y. Ufimtsev, "Method of edge waves in the physical theory of diffraction" translated by U.S. Air Force, Foreign Tech. Div., Wright-Patterson AFB, OH, 1971.
 41. P. Y. Ufimtsev, "Theory of acoustical edge waves", J. Acoust. Soc. Am., vol. 86, pp. 463-474, 1989.
 42. G. Apaydin, L. Sevgi, "Diffraction and fringe effects in EMC problems", in 2017 IV Int. Electromag. Compat. Conf. (EMC Turkiye), Ankara, Turkey, Sept. 24-27, pp. 1-4, 2017.
 43. Y. Z. Umul, "Fringe waves in wedge diffraction", Optik, vol. 123, pp. 217-222, 2012.

44. M. Ando, "Radiation pattern analysis of reflector antennas", *Elect. Comm. In Japan* vol. 68, pp. 93-102, 1985.
45. O. Breinbjerg, "A theoretical examination of the physical theory of diffraction and related equivalent currents", *Electromag. Inst., Tech. Univ. Denmark, Denmark*, 1987.
46. K. M. Mitzner, "Incremental length diffraction coefficients", *Northrop Corp. Aircraft Division Technical Report AFAL-TR-73-296*, 1974.
47. R. A. Shore, A. D. Yaghjian, "Incremental diffraction coefficients for planar surfaces" *IEEE Trans. Ant. Propag.*, vol. 36, pp. 55-70, 1988.
48. T. B. Hansen, R. A. Shore, "Incremental length diffraction coefficients for the shadow boundary of a convex cylinder", *IEEE Trans. Ant. Propag.*, vol. 46, pp. 1458-1466, 1998.
49. A. Michaeli, "Equivalent edge currents for arbitrary aspects of observation", *IEEE Trans. Ant. Propag.*, vol. AP-32, pp. 252-258, 1984.
50. E. F. Knott, "The relationship between Mitzner's ILDC and Michaeli's equivalent currents", *IEEE Trans. Ant. Propag.*, vol. AP-33, pp. 112-113, 1985.
51. Y. Z. Umul, "Modified theory of physical optics", *Opt. Express*, vol. 12, pp. 4959-4972, 2004.
52. W. Burnside, C. Yu, R. Marhefka, "A technique to combine the geometrical theory of diffraction and the moment method", *IEEE Trans. Ant. Propag.*, vol. 23, pp. 551-558, 1975.
53. G. Apaydin, L. Sevgi, P. Y. Ufimtsev, "Fringe integral equations for the 2-D wedges with soft and hard boundaries", *Radio Sci.*, vol. 51, pp. 1570-1578, 2016.
54. G. Apaydin, L. Sevgi, "The two-dimensional nonpenetrable wedge scattering problem and a MATLAB-based fringe wave calculator", *IEEE Ant. Propag. Mag.*, vol. 58, pp. 86-93, 2016.
55. F. Hacivelioglu, L. Sevgi, P. Y. Ufimtsev, "Wedge diffracted waves excited by a Line source: exact and asymptotic forms of fringe waves", *IEEE Trans. Ant. Propag.*, vol. 61, pp. 4705-4712, 2013.

Incremental theory of diffraction

56. G. Pelosi, S. Maci, R. Tiberio, "Incremental length diffraction coefficients for an impedance wedge" *IEEE Trans. Ant Propag.*, vol. 40, pp. 1201-1210, 1992.
57. D. I. Butorin, P. Y. Ufimtsev, "Explicit expressions for an acoustic edge wave

- scattered by an infinitesimal edge element", *Sov. Phys. Acoust.*, vol. 32, pp. 283-287, 1986.
58. P. Y. Ufimtsev, "Elementary edge waves and the physical theory of diffraction", *Electromag.*, vol. 11, pp. 125-159, 1991.
 59. R. Tiberio, S. Maci, "An incremental theory of diffraction: scalar formulation" *IEEE Trans. Ant. Propag.*, vol. 42, pp. 600-612, 1994.
 60. R. Tiberio, S. Maci, A. Toccafondi, F. Capolino, "Applications of the incremental theory diffraction to the scattering by planar configurations", *IEEE AP-S/URSI 1994*, Seattle, WA, June 20-24, vol. 1, pp. 610-613, 1994.
 61. R. Tiberio, S. Maci, A. Toccafondi, "An incremental theory of diffraction electromagnetic formulation", *IEEE Trans. Ant. Propag.*, vol. 43, pp. 87-96, 1995.
 62. R. Tiberio, A. Toccafondi, A. Polemi, S. Maci, "Incremental theory of diffraction: a new-improved formulation", *IEEE Trans. Ant. Propag.*, vol. 52, pp. 2234-2243, 2004.
 63. A. Toccafondi, S. Maci, A. Polemi, M. Albani, F. Capolino, "The contributions of Prof. Roberto Tiberio to the incremental theory of diffraction for electromagnetic", *IEEE Ant. Propag. Soc. Int. Symp.*, Albuquerque, NM, July 9-14, pp. 2461-2464, 2006.
 64. D. Erricolo, M. Albani *et. al.*, "Experimental and theoretical validation for the incremental theory of diffraction", *IEEE Trans. Ant. Propag.*, vol. 56, pp. 2563-2571, 2008.
 65. A. Polemi, G. Carluccio, M. Albani, A. Toccafondi, S. Maci, "Incremental theory of diffraction for complex point source illumination", *Radio Sci.*, vol. 42, pp. 1-13, 2007.
 66. A. Toccafondi, C. D. Giovampaola, M. Albani, D. Erricolo, "ITD formulation for the high-frequency scattering by moderately sized elliptic cylinders", *IEEE Ant. Propag. Soc. Int. Symp.*, San Diego, CA, July 5-11, pp. 1-4, 2008.
 67. G. Carluccio, M. Albani, S. Maci, "Incremental beam diffraction from flat reflectors", in *3rd Euro. Conf. on Ant. Propag.*, Berlin, Germany, March 23-27, pp. 2119-2122, 2009.
 68. S. M. Canta, D. Erricolo, A. Toccafondi, "An ITD formulation for the double diffraction by a pair of wedges illuminated by an EM complex source point expansion", *Proc. Fourth Euro. Conf. Ant. Propag.*, Barcelona, Spain, April 12-16, pp. 1-5, 2010.

Multiple diffraction

69. G. Millington, R. Hewitt, F. S. Immirzi, "Double knife-edge diffraction in field-strength predictions", *Proc. IEE – Part C: Monographs*, vol. 109, pp. 419-429, 1962.

70. J. Deygout, "Multiple knife-edge diffraction of microwaves", IEEE Trans. Ant. Propag., vol. AP-14, pp. 480-489, 1966.
71. R. Tiberio, R.G. Kouyoumjian, "Calculation of high-frequency diffraction by two nearby edges illuminated at grazing incidence" IEEE Trans. Ant. Propag., vol. AP-32, pp. 1186-1196, 1984.
72. C. L. Giovaneli, "An analysis of simplified solutions for multiple knife-edge diffraction" IEEE Trans. Ant. Propag., vol. AP-32, pp. 297-301, 1984.
73. G. Carluccio, F. Puggelli, M. Albani, "Generalisation of UTD Double Diffraction to the case of Impenetrable Wedges with relatively general boundary conditions", IEEE Trans. Ant. Propag., vol. 62, pp. 3829-3834, 2014.
74. J. D. Hunter, R. H. T. Bates, "Secondary diffraction from close edges on perfectly conducting bodies", Int. J. Electron., vol. 32, pp. 321-333, 1972.
75. L. P. Ivrisimtzis, R. J. Marhefka, "A note on double edge diffraction for parallel wedges", IEEE Trans. Ant. Propag., vol. 39, pp. 1532-1537, 1991.
76. F. Capolino, S. Maci, "Uniform high-frequency description of singly, doubly and vertex diffracted rays for a plane angular sector", J. Electromag. Waves Appl., vol. 10, pp. 1175-1197, 1996.
77. M. Schneider, R. Luebbers, "A general, uniform double wedge diffraction coefficient", IEEE Trans. Ant. Propag., vol. 39, pp. 8-14, 1991.
78. L. P. Ivrisimtzis, R. J. Marhefka, "Double diffraction at a coplanar skewed edge configuration", Radio Sci., vol. 26, pp. 821-830, 1991.
79. F. Capolino, M. Albani, S. Maci, R. Tiberio, "Double diffraction at a pair of coplanar skew edges", IEEE Trans. Ant. and Propag., vol. 45, pp. 1219-1226, 1997.
80. M. Albani, "A uniform double diffraction coefficient for a pair of wedges in arbitrary configuration", IEEE Trans. Ant. Propag., vol. 53, pp. 702-710, 2005.
81. P. D. Holm, "UTD diffraction coefficients for higher order wedge diffracted fields", IEEE Trans. Ant. Propag., vol. 44, pp. 879-888, 1996.
82. G. Carluccio, F. Puggelli, M. Albani, "A UTD triple diffraction coefficient for straight wedges in arbitrary configuration", IEEE Trans. Ant. Propag., vol. 60, pp. 5809-5817, 2012.

4. Diffraction from an Isolated Structure

This section uses numerical modelling of isolated structures to validate the diffraction theories discussed in the previous chapter. We begin by reviewing literature comparisons between various analytical theories and with different numerical models. Based on these reviews, the geometrical theory of diffraction (GTD) is chosen as the most suitable analytic model, despite its approximate nature. We then illustrate the process of finite element modelling using the simulation tool COMSOL. Numerical results are presented to demonstrate the cylindrical wave nature of the diffracted field and the diffraction pattern from a wedge, initially for TE polarization. Comparison between numerical and analytical results is made for diffraction from a wedge at different incident angles. Numerical and analytical diffraction coefficients for a mesa and a half-plane are then compared, to justify the selection of the GTD wedge model as the basis of a simplified scattering model in Chapter 7. Finally, we repeat the main steps for TM polarization.

4.1 Analytical model

The diffraction theories discussed in the previous chapter were the geometrical theory of diffraction (GTD) [1, 2], the uniform theory of diffraction (UTD) [3], the uniform asymptotic theory (UAT) [4], the physical theory of diffraction (PTD) [5] and incremental theory of diffraction (ITD) [6]. To study the diffraction from the complex structures expected in NEMS devices, a benchmark analytical model is needed to provide a simple but accurate approximation to the exact solution of the scattering problem.

The two main diffraction theories, GTD and PTD, have always been discussed together and compared since they were established [7, 8]. The ray approach in GTD is simple and the field expression is straightforward for all scenarios. In PTD, since the induced surface currents are integrated for the scattered field, the field components are finite in all regions. For the far field of a simple structures (such as a strip), both geometrical and physical approaches provide close approximations to the exact solution of the scattering problem. The asymptotic expansion of the GTD is always correct [9], while PTD yields good numerical estimates with an erroneous asymptotic expansion [10].

In PTD, the total electric field expression is:

$$E_t = E_i + \iint_{\Sigma_{1,2}} \bar{\bar{G}} J_0 ds + \iint_{\Sigma_{1,2}} \bar{\bar{G}} J_1 ds \quad (4.1)$$

Here $\bar{\bar{G}}$ represents the standard free-space dyadic Green's function. $\Sigma_{1,2}$ and s are the two surfaces of the structure (see Figure 3.12) and their unit area respectively. The first term is the incident field, and the remaining terms are the scattered field produced by the uniform source $J_{0,1}$ radiating in the free space respectively, formulated from the Hertz potential [11-13]. The integrand involving J_0 is also referred as the physical optics integral. Unlike GTD, PTD does not provide an explicit field solution and the physical optics integral must be evaluated asymptotically. This integral can sometimes be difficult to compute for a complex structure [14].

For near field scattering problem, GTD exhibits significant differences from the exact solution in the caustic and transition region. The uniform theories of GTD improved the approximation and yield high-frequency asymptotic solutions of the edge diffraction problem that are uniformly valid in the entire space. In general, the uniform asymptotic expansion of the exact field solution can be carried out by two procedures: the Van der Waerden [15] (UAT); and the Pauli [16] and Clemmow [17] (UTD) methods. The field expression in UAT and UTD are different as they employ different assumptions about the solutions (ansatz). The UAT solution can be extended systemically to higher order in k^{-1} (k is the wavenumber) but this cannot be done in UTD. Also, the UAT solution has continuous derivatives everywhere, but UTD solution becomes discontinuous at the shadow boundaries [18-20].

Numerical computation is often used to provide accurate solutions to complex EM wave, particularly to calculate scattered fields [21-24]. Common methods are the finite element method (FEM), the method of moments (MoM) and the finite difference time domain method (FDTD) [25-28]. In general, numerical results from all the above methods agree with the exact solution of the diffraction problem and are often regarded as references. For example, Ozgun presented the fringe field obtained from FEM for differently sized wedges and mesas [29, 30]. The results were also compared to the MoM and good agreement was obtained. Apaydin and Hacivelioglu *et al.* investigated the total and fringe field obtained using PTD for wedges numerically and found the relative error to be 3% or less compared to the exact solution [31, 32]. The diffracted fields from wedges obtained

using UTD were compared with MoM and FDTD and good agreement was again obtained [28, 33, 34]. A MATLAB-based simulation tool was also developed to compare the scattering from wedges obtained using high frequency asymptotics and the FDTD method. The numerical results proved that the diffraction pattern calculated from PTD and UTD are approximated accurately for wedges of any size [35, 36].

In optical imaging, scattering is a far field problem. However, multiple diffraction from closely spaced structures like slots or mesa arrays often arises in NEMS devices. Therefore, PTD is not suitable as the computation of the physical optics integral (Equation 4.1) may often be too difficult. UTD is also less preferable due to the complexity of the calculation of the diffraction coefficients compared to GTD, which is sufficiently accurate in the far field. Therefore, GTD is chosen as the benchmark theory to compare with the numerical results to study the optical imaging of NEMS.

4.2 Finite element modelling

FEM is a popular numerical method for solving complex problems in different physical disciplines. It defines a domain of interest using tetrahedral elements that are subject to specified boundary conditions [37-41]. Here, its advantage is the ability to solve general diffraction problems without complex mathematics. However, its scope and accuracy are proportional to the available computational resources, and this has severely restricted the research that has been carried out using it [42-45].

Here the “Wave Optics” module of the commercial package COMSOL® was used [46]. The general steps for FEM are first to design a layout for simulation. Space is then discretised into a set of polygonal or polyhedral mesh elements, usually of the size of a fifth of the incident wavelength. The mesh nodes are concentrated at corners where field variation is the largest, using a mesh optimisation algorithm. A solution to the wave equation is then obtained by solving compatibility equations at each node and boundary conditions at the domain boundaries.

A right-angled corner model was created for the comparison with analytical results obtained from GTD, previously presented in Figure 3.6. Figure 4.1 shows the simulation model, with a square box of length L defining the simulation region. Two inner squares

form a perfectly matched layer (PML) and an absorber layer, to remove unwanted propagating field components. The absorber removes majority of the cylindrical wave interference caused by the PML boundary at parallel (or almost parallel) incidence. The wedge is modelled as a perfect electric conductor (PEC) with side length of $L/2$ shown in grey with PEC boundary conditions shown as red lines. With part of the wedge lying inside the PML and absorber, the section of the wedge in the air region has a side length $\Lambda/2$. This is the only contributor to scattering.

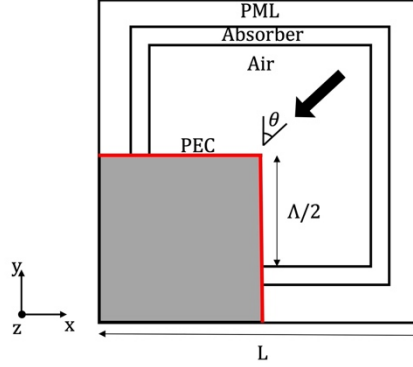


Figure 4.1: Simulation model of an isolated right-angled wedge.

The x- and y-directions represent the horizontal and vertical axis, and the z-direction is the out-of-page axis. The only non-zero component of the electric field for TE incidence and for the magnetic field for TM incidence are the z-components, E_z and H_z respectively. The corresponding x- and y-components of the EM wave are $H_{x,y}$ (TE) and $E_{x,y}$ (TM). The illumination is in a form of a background wave in the whole simulation space at an incident angle θ from the y-axis.

For a wedge diffraction, the angle of illumination is generally separated into two types: $\theta = 0^\circ$ to $+90^\circ$ (two-side illumination) and $\theta = 0^\circ$ to -90° (one-side illumination). Here we refer them as the leading and trailing edge diffraction. The major difference lies in the field components. Figure 4.2 shows the subspaces of the diffraction for a) leading and b) trailing edge illumination. The black lines indicate the shadow boundaries that divide the space into different regions. For leading edge diffraction, the incident wave is reflected by both the top and side surfaces of the wedge in region 1 and 3 respectively. The incident and diffracted waves propagate in the whole space in two-side illumination. For trailing edge illumination, the incident field only propagates in region 1 and 2, and is

partly reflected by the top surface of the wedge in region 1. However, the diffracted field from the corner still propagates in all three regions.

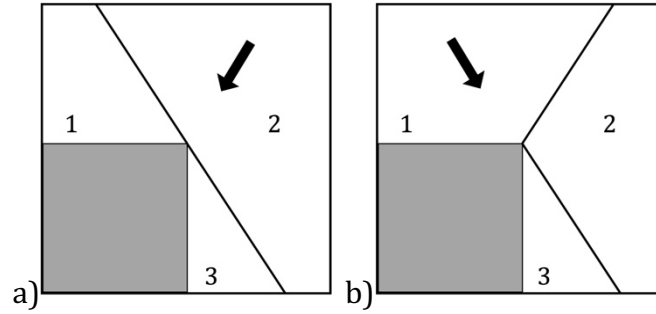


Figure 4.2: Subspaces of a) leading and b) trailing edge diffraction.

Clearly, the simulation model in Figure 4.1 must be adjusted for the two cases. Part of the absorber layer must be “turned off”, as it will prevent the incident wave from travelling into the air domain. However, most of the absorber layer must remain “on” to remove interference. To remove additional calculations for the region, the model in Figure 4.1 must be altered to calculate the reflected field components only. Figure 4.3 shows the modified geometry for the reflected fields from a) the top and b) the side surfaces. The required modification is to extend the corresponding surfaces across the whole model to obtain the reflected field at different angles of incidence (indicated by the red lines). This ensures the total field contains only the reflected and incident fields.

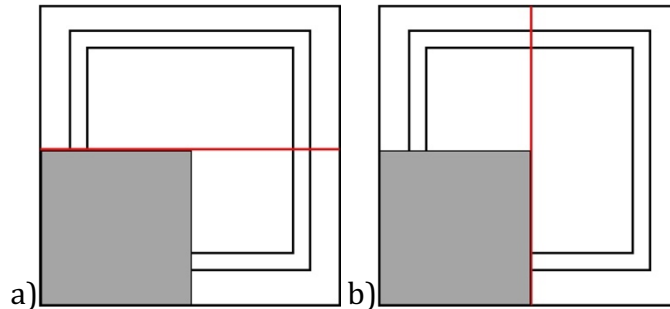


Figure 4.3: Modified model used to calculate reflected field components from a) top and b) side PEC boundaries.

To compare the numerical results to the analytical results from GTD, the simulated electric fields are exported from COMSOL to MATLAB for further processing. Here we follow Ozgun [29] for the scattered field calculation in TE polarization. The total field is calculated from the model in Figure 4.1, while the incident and reflected fields are obtained from the models in Figure 4.1 and Figure 4.3. The area of the regions (in Figure

4.2) at each angle of incidence is calculated by comparing the x- and y-coordinates with the shadow boundaries in MATLAB. The diffracted field E_d can then be calculated from subtracting the incident E_i and reflected E_r fields from the total field E_t considering the field components in each region $R_{1,2,3}$, expressed as:

$$E_{d_{LEAD}} = E_t - R_1 E_{r_t} - R_2 E_i - R_3 E_{r_s} \quad E_{d_{TRAIL}} = E_t - R_1 E_{r_t} - R_2 E_i \quad (4.2)$$

Here $E_{d_{LEAD}}$ and $E_{d_{TRAIL}}$ denotes the diffracted fields of a leading and trailing edge and E_{r_t, r_s} represents the top and side reflections. Similarly, diffracted magnetic field H_d can be calculated in TM polarization with the same method, by replacing the H field with the E field in the above equations.

Once the diffracted field is obtained, the diffraction coefficient can be calculated from the cylindrical wave expression. Recall the diffracted field expression from Equation 2.29:

$$E_d = \frac{E_0 D e^{-jkr}}{\sqrt{r}} \quad (4.3)$$

Assuming the magnitude of the incident field is unity, the diffraction coefficient relies heavily on the radial distance r from the origin, which is the corner of the wedge. If a small unit circle is chosen ($r < 1\mu m$), there are not enough data points to provide the diffraction pattern. However, if r is too big, the interference from the PML will become significant and affect the accuracy of the result (especially for TM polarization). Also, since the diffraction coefficient D is complex, the diffraction pattern is compared with the parameter D_p , expressed as:

$$D_p = D \sqrt{2\pi k} e^{+j\frac{\pi}{4}} \quad (4.4)$$

Where D_p is equivalent to the trigonometric part of the diffraction coefficient of Keller's expression in GTD in Equation 2.31.

4.3 Diffraction from a wedge

In this section, TE results of an isolated right-angled corner are presented. First, the cylindrical nature of the diffracted field is verified using FEM. Then, the diffraction coefficients are reviewed at each angle of incidence and at different radii. Next, the diffraction coefficients are compared with the analytical result from GTD. Finally, the diffracted fields of a mesa and a half-plane are compared with the shifted sum of the wedge diffraction. The size of the simulation model is $85\mu m$ as it is the largest possible

size with our available resources. The simulation wavelength is $0.5\mu m$ and the range of the angle of incidence θ is -90° to $+90^\circ$ at step size of 10° .

4.3.1 The nature of cylindrical wave

The diffracted field is calculated from Equation 4.2 and considered in terms of its magnitude and real component. Figure 4.4 shows the diffracted field from a right-angled corner at incident angle $\theta = 30^\circ$, indicated by the red arrows. The white box in the contour plots is the wedge corner. The left panel shows the magnitude of the diffracted field $|E_d|$ in an expanded region from -5λ to $+5\lambda$. The shadow boundaries are highlighted as cyan/light green in the contour plot due to the larger magnitude. The field strength is weaker towards the surfaces of the wedge corner. The middle panel shows the real component of the diffracted field $Re(E_d)$ in the same area. The wave has a circular pattern with a magnitude decaying from the origin. There is also a sign change of the real component across the shadow boundaries. The right panel shows the magnitude of the diffracted field in a polar plot at $r = 4\mu m$. It shows the magnitude of the diffracted field at the shadow boundaries is approximately 5 times of that in the incidence region.

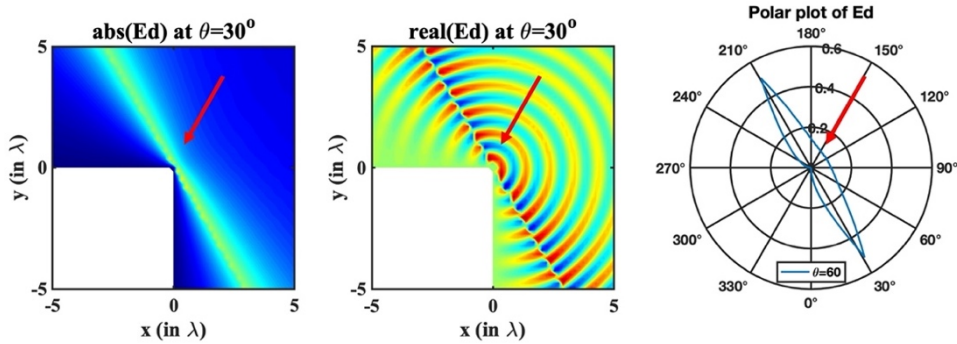


Figure 4.4: Diffracted field from a wedge corner at $\theta = 30^\circ$, magnitude (left); real component (middle); and magnitude in polar plot at $r = 4\mu m$ (right).

Figure 4.4 implies that the wavefront of the simulated diffracted field is circular and centred on the wedge corner. Now we take the magnitude of the scattered field along the x- and y-axis and compare it with a cylindrical wave. Figure 4.5 compares the real and imaginary component of the scattered field from the corner at $\theta = 30^\circ$ along the x-axis (left) and the y-axis (right) with a complex cylindrical wave with a $\pi/4$ phase difference. The left panel shows that the real and imaginary parts (black and green lines) of the diffracted field along the x-axis decay as a cylindrical wave (red and blue dashed lines), except for $x < 0.5\lambda$. Similarly, the right panel shows that scattered field along the y-axis

also decays as a cylindrical wave. Consequently, it can be concluded that the diffracted field from a wedge is indeed a cylindrical wave.

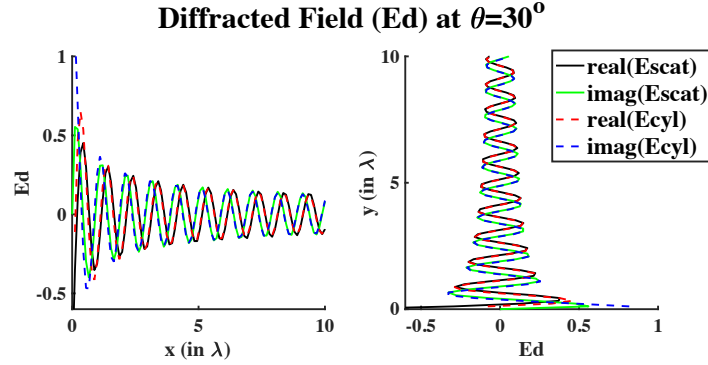


Figure 4.5: Comparison between the diffracted field from a wedge corner at $\theta = 30^\circ$ with a cylindrical wave, along the x-axis (left) and y-axis (right), for TE incidence.

4.3.2 Diffraction coefficients

Figure 4.6 shows the polar plots of the diffraction pattern magnitude D_p from a wedge corner at angles of incidence varying from $\theta = 90^\circ$ to -60° at radius $r = 4\mu m$. The red arrows show the directions of the incident wave. The diffraction patterns vary with the angles of incidence with an approximately identical magnitude for two-side illumination and grazing incidences ($\theta = 90^\circ$ and 0°). For one-side illumination, the diffraction patterns between the shadow boundaries superimpose so the magnitude of the diffraction coefficient is larger and different to the other incidences.

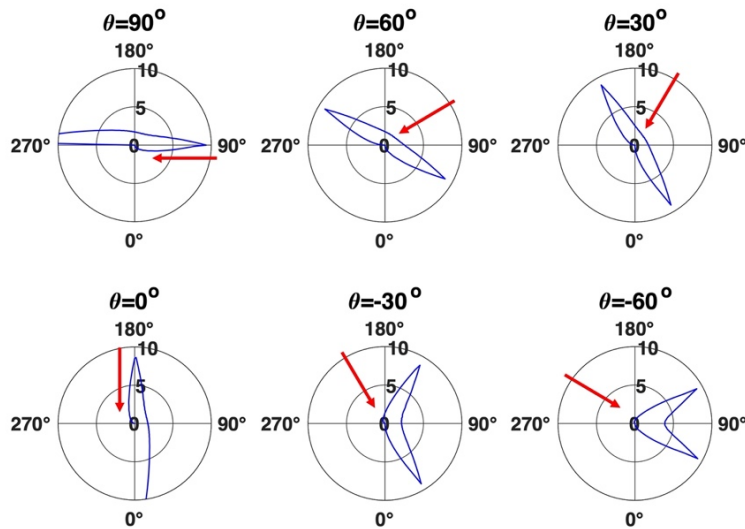


Figure 4.6: Diffraction pattern D_p of a wedge diffraction at incident angles from $\theta = 90^\circ$ to -60° at a step of 30° , for TE incidence.

Figure 4.7 shows the magnitude of the leading and trailing edge diffraction patterns at $\theta = 30^\circ$, in blue and red respectively. The diffraction patterns at the shadow boundary between regions 2 and 3 are almost identical, but they are different in region 1 and 2, leading to a different angular variation. In particular, the values of D_p at diffraction angle $\phi = 90^\circ$ ($\phi = 0^\circ$ at the vertical surface of the wedge) for trailing edge diffraction is larger. At $\phi = 180^\circ$, the leading-edge pattern is double of that from the trailing edge. These two diffraction angles are important when considering multiple diffraction and spatial filtering by a lens during simulation of optical imaging and will now be defined as D_x ($\phi = 90^\circ$) and D_y ($\phi = 180^\circ$). We therefore consider their variation in more detail.

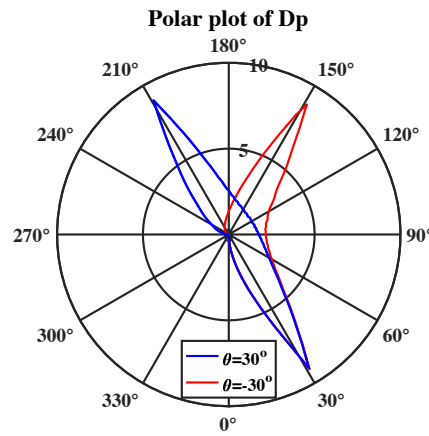


Figure 4.7: Leading and trailing edge diffraction pattern, for TE incidence at $\theta = 30^\circ$.

We first note that a trend can be observed in the angular variations of D_x and D_y . Figure 4.8 shows the magnitude of the diffraction coefficient D_p at $\phi = 90^\circ$ (D_x represented in blue) and $\phi = 180^\circ$ (D_y represented in red) calculated from FEM (dots) and GTD (solid lines) for a) two-side and b) one-side illumination. Figure 4.8a shows a symmetrical trend for D_x and D_y for two-side illumination at $\theta = 45^\circ$ because the wedge is right-angled. Figure 4.8b shows that the vertical diffraction is positive and decreasing asymptotically towards zero at larger incident angles, while the horizontal diffraction increases with incidence angle. The diffraction coefficient at $\theta = -90^\circ$ from FEM is an anomaly, as the incidence is parallel to the wedge and no diffraction occurs. In both illuminations, the numerical results closely follow the analytical values obtained from GTD at $r = 4\mu m$.

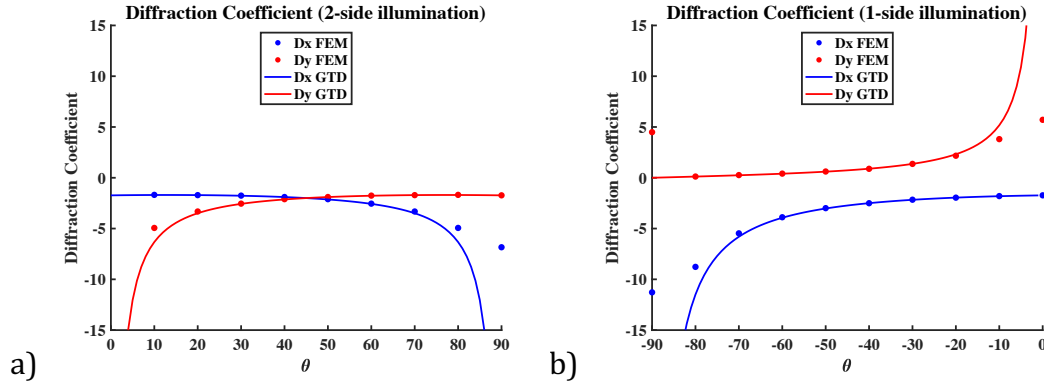


Figure 4.8: Diffraction pattern D_x and D_y obtained from FEM and GTD for a) leading and b) trailing edge diffractions, for TE incidence.

If we define the vertical diffraction coefficient of leading and trailing edges as D_{y1} and D_{y2} respectively, the ratio D_{y2}/D_{y1} defines the relative significance of the two. Figure 4.9 shows the analytical (red) and numerical (blue) values of D_{y2}/D_{y1} at all angles. Again, the numerical values closely follow the analytical results. The larger the value of θ , the weaker the diffraction of a trailing edge compared to the leading edge. This is useful information for constructing the imaging model and will be discussed in more detail later.

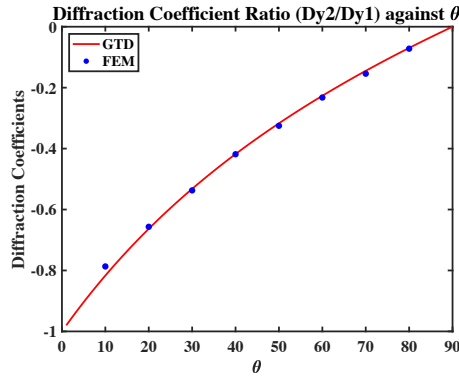


Figure 4.9: Ratio of diffraction coefficient D_{y2}/D_{y1} at $\theta = 10^\circ$ to 90° , for TE incidence.

In reality, numerical values of the diffraction coefficient depend on the observation radius r . Figure 4.5 has already shown a discrepancy between the numerical results and a cylindrical wave for $r < 0.5\lambda$. We therefore now consider the variation of the diffraction coefficient with r in more detail. Figure 4.10 shows the diffraction coefficients D_y for leading (red) and trailing (green) edge diffraction from both numerical and analytical methods for r ranging from $0.1\mu m$ to $4\mu m$. Since the analytical expression from GTD is independent of r , these values are constant, and shown as black and grey lines. The

average values of the FEM results are shown as coloured dashed lines. For $r < 1\mu m$, the magnitude of both diffraction coefficients is smaller than the analytic values and the limits. However, as r increases, both numerical diffraction coefficients gradually tend to the analytical values, implying that GTD and FEM results agree in the far field.

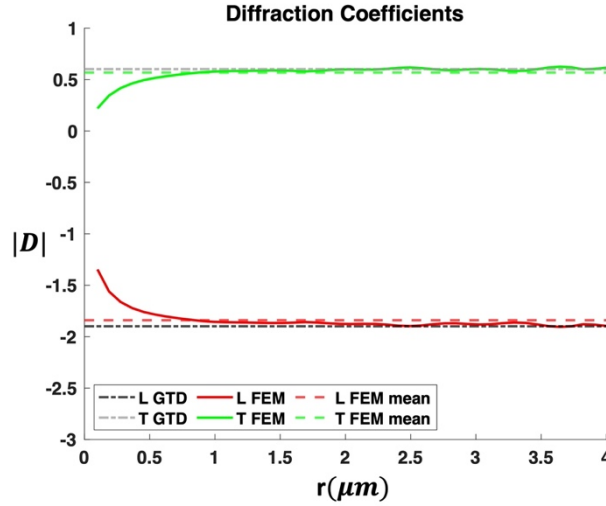


Figure 4.10: Diffraction coefficients for leading (red) and trailing (green) edges at different r , for TE polarization.

Now we consider the variation of the diffraction coefficient with the diffraction angle ϕ . Figure 4.11 shows the diffraction pattern for a right-angled wedge diffraction at $\theta = 30^\circ$ obtained from FEM (blue) and GTD (red). The magnitudes of the diffraction coefficient from both methods are almost identical, except for the transition regions around the shadow boundaries. In the transition regions, diffraction coefficients obtained using GTD is Infinite, due to the asymptotic problems mentioned in the previous chapter.

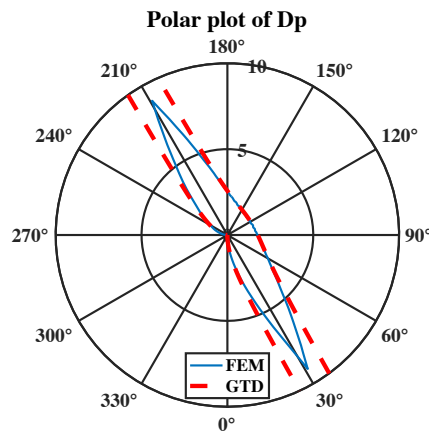


Figure 4.11: Wedge diffraction pattern at $\theta = 30^\circ$ from FEM (blue) and GTD (red), for TE polarization.

4.4 Diffraction from a mesa and a half-plane

We have so far compared diffraction from a right-angled wedge in two-side and one-side illumination using numerical and analytical results. For an isolated mesa, diffraction can be considered as the sum of leading and trailing edge diffraction. Figure 4.12 shows the diffracted field of an isolated mesa of width $w = 2\mu\text{m}$ in a simulation space of $85\mu\text{m}$ at $\theta = 30^\circ$. The absorber layer is removed in this model as it is complicated to modify it to remove only the unnecessary field component. The left and middle panels show the magnitude and the real component of the diffracted field in an expanded view from -10λ to $+10\lambda$ ($\lambda = 0.5\mu\text{m}$). The shadow boundaries from the leading and trailing edge diffraction are shown in cyan. However, since the shadow boundaries of the top surface reflection from the two corners are close, the diffracted field in between is also superimposed and amplified. The right panel shows a polar plot of the magnitude of the diffracted field. The angles for the shadow boundaries deviate from $\theta = 30^\circ$, as the angular field strength is measured at $r = 4\mu\text{m}$ from the point (0,0), instead of from the corner of the wedge. Because of this shift in origin, the analytical diffraction coefficient is more complicated to compute.

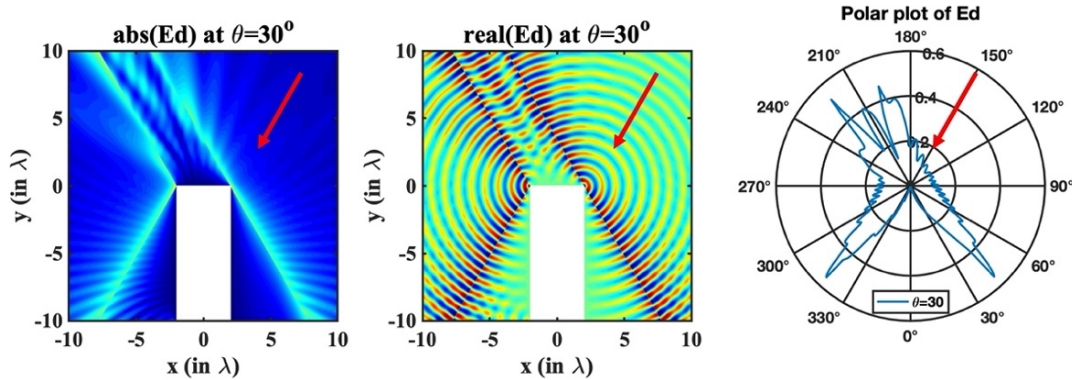


Figure 4.12: Diffracted field from a mesa at $\theta = 30^\circ$, magnitude (left); real component (middle); and magnitude in polar plot at $r = 4\mu\text{m}$ (right), for TE polarization.

Figure 4.13 shows the diffracted field from wedge at a) $\theta = 30^\circ$ and b) $\theta = -30^\circ$ in the same simulation environment as Figure 4.12. Since the absorber is removed, Figure 4.13a shows significant interference around the transition region compared to the left panel in Figure 4.4 at the same angle of incidence. Figures 4.13c and d show diffracted fields from an isolated mesa and a half-plane obtained from a shifted sum of the numerical results of the separate wedge diffractions in Figures 4.13a and b. The diffracted fields in Figure 4.13c and d clearly show excellent resemblance to the numerical results of the

corresponding structure (Figure 4.12 left panel and Figure 4.14 left panel). This conclusion implies the wedge diffraction model can closely approximate diffraction from a mesa and a half-plane simply by summation of the diffraction of two adjacent wedges.

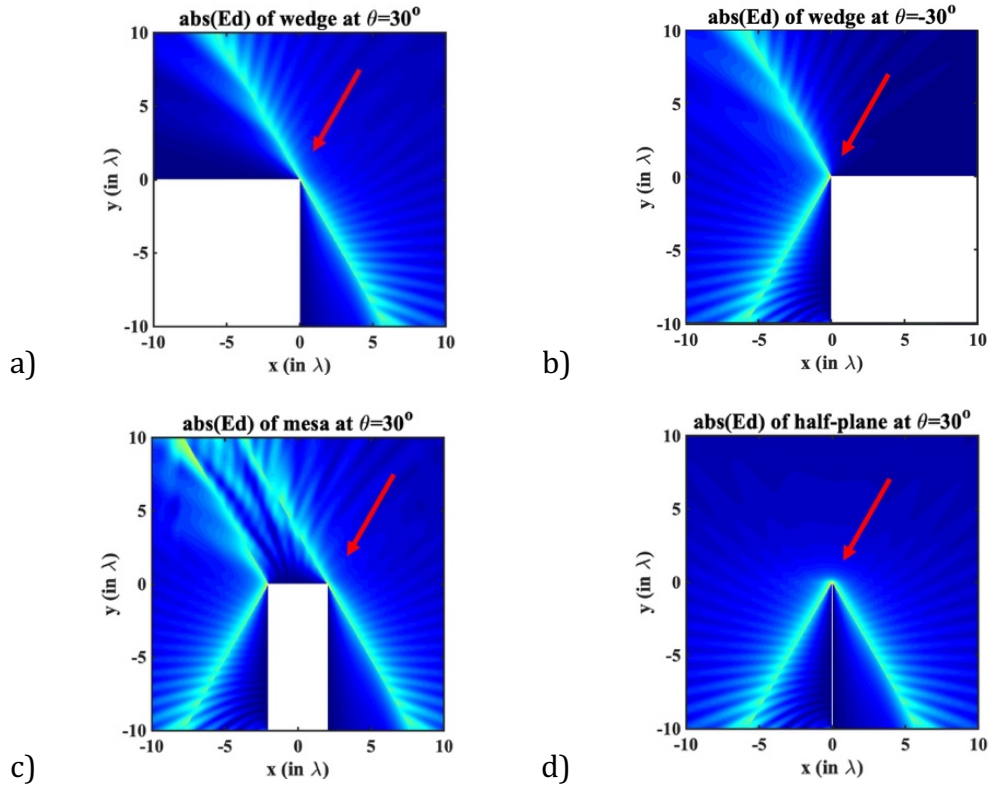


Figure 4.13: Diffracted field from a wedge at a) 2-side illumination and b) 1-side illumination, for TE polarization. Shifted sum obtained for c) a mesa and d) a half-plane.

Now we simulate diffraction from a half-plane, by approximating it as a very narrow mesa. Figure 4.14 shows the diffracted field from an isolated mesa of width $w = 0.1\mu\text{m}$ at $\theta = 30^\circ$ in an expanded view of 20λ span. The left and middle panels show the magnitude and the real component of the diffracted field. The wave front of the diffracted field is almost perfectly circular, with a thin field line of large amplitude at $\phi = 210^\circ$ that represents the superposition between the shadow boundaries. The right panel shows the comparison of the analytical (red) and numerical (blue) diffraction coefficients. The analytical values approach infinity in the transition region (from Equation 2.30) when the numerical value is finite in the same region. This result shows that half-planes can be accurately modelled using simple summation of GTD wedge diffraction effects.

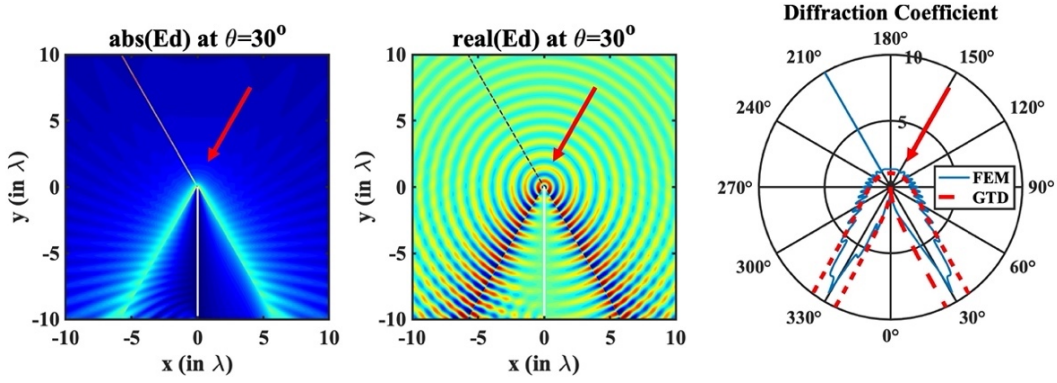


Figure 4.14: Diffracted field from a half-plane at $\theta = 30^\circ$, magnitude (left); real component (middle); and the diffraction coefficient, compared between FEM (blue) and GTD (red) at $r = 4\mu\text{m}$ (right), for TE polarization.

We now consider whether the reverse can be achieved, namely whether a mesa can be modelled by summation of GTD half-plane effects. For example, Figure 4.15 shows the sum of the numerical diffracted field of two adjacent half-planes at a distance of $2\mu\text{m}$. Compared to the numerical result of a PEC-bounded mesa in Figure 4.12, the superposition of the field above the mesa is missing. This is due to the absence of the top shadow boundary for a half-plane. To summarise, the diffraction pattern of a half-plane and an isolated mesa can both be estimated by summing the separate effects of the leading and trailing edge. However, the diffraction of a mesa cannot be recreated by two half-planes. Therefore, the wedge model of the Keller's expression is selected for the imaging model developed later on.

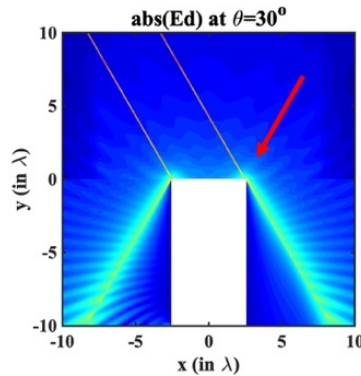


Figure 4.15: Shifted sum of two adjacent half-planes at distance of $2.5\mu\text{m}$.

4.5 TM Polarization

In the previous section, the diffraction of the electric field from a wedge and mesas are discussed in TE polarization. For TM polarization, the diffraction of EM waves is very similar. TM results are shown in this section from the same simulation model used previously in TE results. The calculation in MATLAB is carried out on the H field and the results are plotted in terms of the electric field to simplify comparison with the TE case.

Figure 4.16 shows the diffracted field from a wedge of 90° at an incident angle $\theta = 30^\circ$, indicated by the arrows. The left and middle panel shows the magnitude $|E_d|$ and the real component $Re(E_d)$ of the diffracted field in the same region. The patterns are very similar to the TE results in Figure 4.4. implying that the diffracted field is also cylindrical for TM polarization. The right panel shows the diffraction pattern of the corner at a radial distance $r = 4\mu m$ from the corner. The pattern is again similar to the TE case. However, a major difference is that the diffracted field strength for TM polarization is stronger in the shadow/reflection region, and weaker in the incidence region.

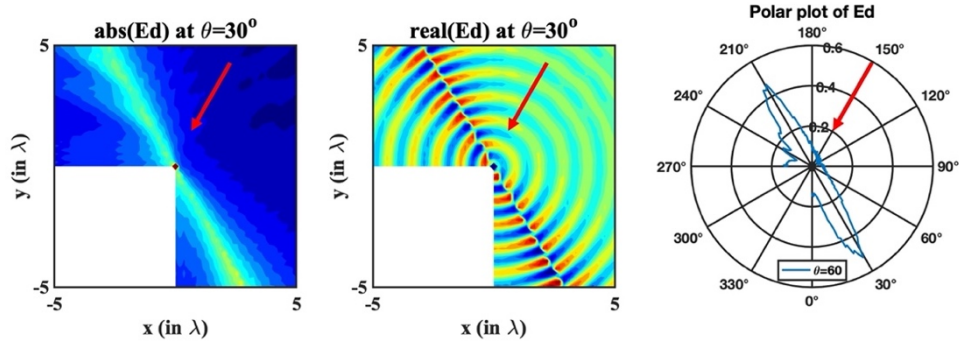


Figure 4.16: Diffracted electric field from a right-angled wedge at $\theta = 30^\circ$, magnitude (left); real component (middle); and magnitude in polar plot at $r = 4\mu m$ (right), for TM polarization.

The diffraction patterns D_p are similar to the TE results in Figure 4.6 and are therefore omitted here. The GTD diffraction coefficients can be obtained for TM polarization (Equation 3.31). Figure 4.17 compares the magnitude of the diffraction coefficient D_p at diffracted angles $\phi = 90^\circ$ (blue, as D_x) and 180° (red, as D_y) with the numerical (FEM, dots) and analytical (GTD, solid lines) results. Figure 4.17a shows their variations for 2-side illumination (leading edge) and Figure 4.17b for a 1-side illumination. For 2-side illumination, the trend is symmetrical with positive coefficients, as it was for the TE

results. However, the TM diffraction coefficients in 1-side illumination vary differently to TE. The magnitude of D_y decreases asymptotically towards zero at larger incident angles and D_x remains approximately constant around zero. For the horizontal diffraction D_x at $\theta = 90^\circ$, an anomaly occurs, as there is no diffraction for incidence parallel to the surface. In general, the results of TM polarization from FEM and GTD agree as expected.

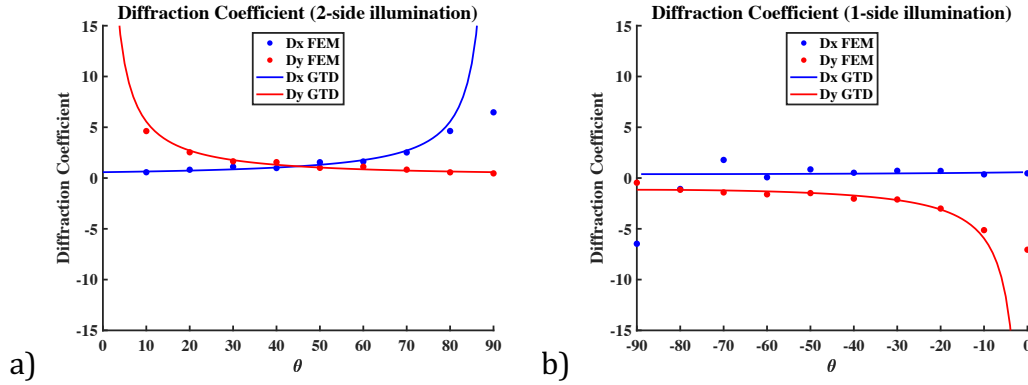


Figure 4.17: Diffraction pattern D_x and D_y comparison between FEM and GTD for a) leading and b) trailing edge diffractions, for TM polarization.

As previously defined, the coefficients D_{y1} and D_{y2} represents the vertical diffraction coefficient for leading and trailing edges, respectively. Figure 4.18 shows the analytical (red line) and numerical (blue dots) variations of their ratio D_{y2}/D_{y1} between $\theta = 0^\circ$ and 90° . In contrast to the TE results, the numerical ratio does not perfectly follow to the analytical values. One possible reason is that interference from the PML layer at larger incident angles is more significant for TM results. However, a general trend is seen in Figure 4.18, namely that the trailing edge is more significant than the leading edge for TM polarization and the significance increases with incident angle.

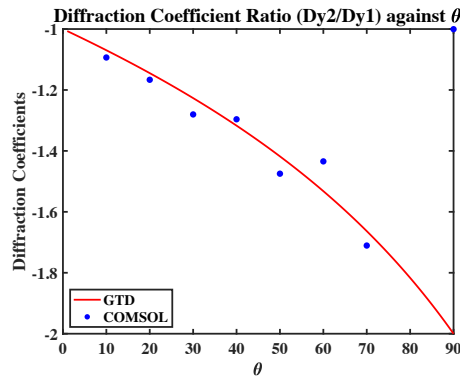


Figure 4.18: Ratio of diffraction coefficient D_{y2}/D_{y1} at $\theta = 10^\circ$ to 90° , for TM polarization.

Next, we review the variation of the diffraction coefficient with the diffraction angle ϕ . Figure 4.19 shows the TM diffraction pattern from a right-angled wedge at $\theta = 30^\circ$ obtained from FEM (blue) and GTD (red). The blue solid line overlaps with the red dashed line for most diffracted angles, except in the transition regions. Compared with the corresponding TE diffraction pattern, the diffraction coefficients around the wedge surface are larger for TM polarization, while the horizontal diffraction is much weaker.

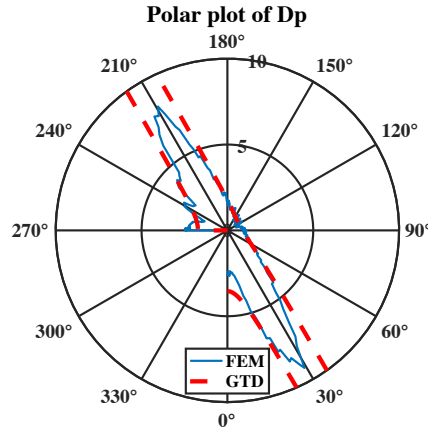


Figure 4.19: Wedge diffraction pattern at $\theta = 30^\circ$ from FEM (blue) and GTD (red), for TM polarization.

We now simulate diffraction from a mesa and a half-plane in TM polarization and compare with the results obtained from wedges. Figure 4.20 shows the diffracted field and the diffraction coefficients of a) a mesa of $w = 2\mu m$ and b) a half-plane (approximated as a narrow mesa of width $w = 0.1\mu m$) in the same simulation environment as Figures 4.12 and Figure 4.14. For TM polarization, diffraction from a mesa is significantly different to the earlier TE results. For example, the electric field clearly propagates along the surfaces of the mesa, implying that secondary diffraction from neighbouring corners is likely to be important in TM polarization. The field inside the shadow and reflection regions are much stronger than in the TE case. However, if we review the field at $r = 4\mu m$, the magnitude and pattern of the diffraction coefficients in both polarizations are similar. For a half-plane, the diffracted field is similar to the TE case; the diffracted field in the transition regions is much stronger and the diffraction pattern matches with the analytical values from GTD.

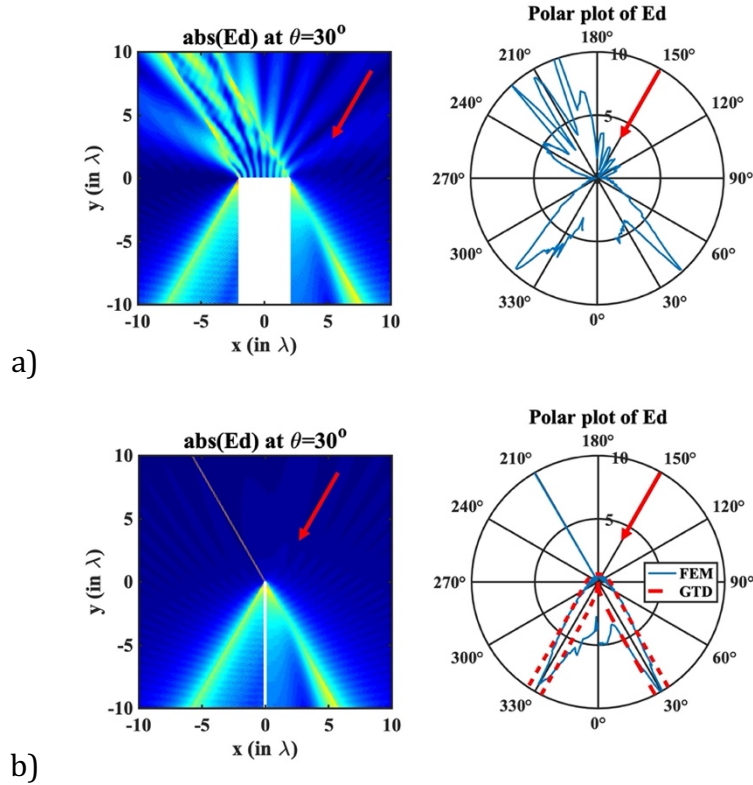


Figure 4.20: Diffracted field from a) a mesa and b) a half-plane at $\theta = 30^\circ$. Left: magnitude; Right: diffraction coefficient in polar plot at $r = 4\mu\text{m}$, for TM polarization.

Finally, we repeat the steps of summing the diffracted fields of a leading and a trailing edge to obtain an approximation to the diffracted field of a mesa and a half-plane for TM polarization. Figure 4.21 shows the diffracted field from a wedge at a) $\theta = 30^\circ$ and b) $\theta = -30^\circ$ and their shifted sum to mimic the diffracted fields of c) a mesa of $w = 2\mu\text{m}$ and d) a half-plane. Since the diffracted field around the wedge surface is now much stronger, secondary diffraction between the two corners is more significant than it was in the TE cases. Since this secondary diffraction is missing in a simple shifted sum, this approximation is much less accurate than might be suggested by the single wedge example in Figure 4.20a. However, since the secondary diffraction from a half-plane is negligible, the approximation is much closer to the result in Figure 4.20b. These conclusions suggest that diffraction from more general structures can again be modelled in terms of GTD coefficients in the TM case, merely provided a sufficient number of scattering terms is taken into account.

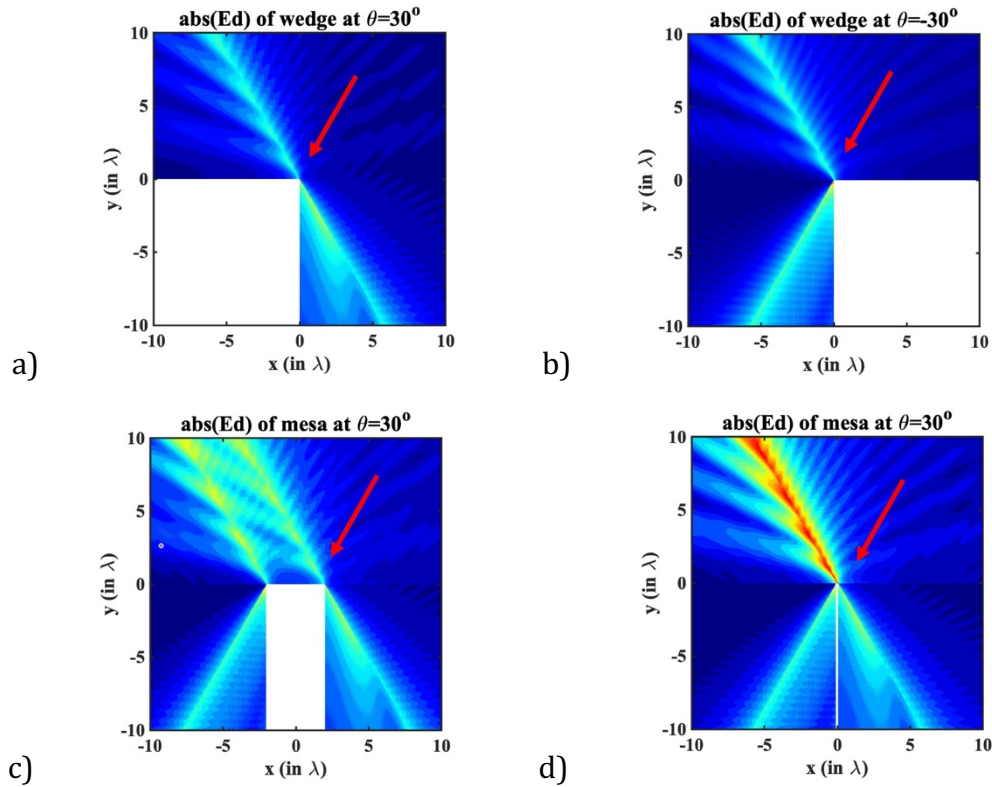


Figure 4.21: Diffracted field from a wedge at a) 2-side illumination and b) 1-side illumination. Shifted sum from (a) and (b) for c) a mesa and d) a half-plane for TM polarization.

4.6 Summary

Analytical results from the GTD approximation have been compared with the full-wave simulation obtained using finite element modelling. Good agreement has been obtained for most of the diffraction scenarios, except in the near field transition regions. For TE polarization, it has been demonstrated that scattering from an isolated mesa can be modelled as a combination of leading and trailing edge diffraction, even for very narrow features. However, for TM polarization, the secondary diffraction must be included to improve accuracy for the mesa case. This conclusion allows the GTD wedge coefficients to be used for the development of a novel model for imaging later. In the next chapter, a further wave simulation method will be compared with the analytical results, namely the modal diffraction model. This approach requires the mesa model to be periodic but removes the PML interference and allows more general changes to the geometry.

References

Literature Review

1. J. B. Keller, "Diffraction by an aperture", J. Appl. Phys., vol. 28, pp. 426-444, 1957.
2. J. B. Keller, "Geometrical theory of diffraction" J. Opt. Soc. Am., vol. 52, pp. 116-130, 1962.
3. R. G. Kouyoumjian, P. H. Pathak, "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface" Proc. IEEE, vol. 62, pp. 1448-1461, 1974.
4. R. M. Lewis, J. Boersma, "Uniform asymptotic theory of edge diffraction" J. Math. Phys., vol. 10, pp. 2291-2305, 1969.
5. P. Y. Ufimtsev, "Approximate computation of the diffraction of plane electromagnetic waves at certain metal bodies" Sov. Phys. Tech. Phys., vol. 27, pp. 1708-1718, 1957.
6. R. Tiberio, S. Maci, "An incremental theory of diffraction: scalar formulation" IEEE Trans. Ant. Propag., vol. 42, pp. 600-612, 1994.
7. P. Beckmann, "The depolarization of electromagnetic waves", Golem Press, Colorado, 1968.
8. T. Senior, P. Uslenghi, "Comparison between Keller's and Ufimtsev's theories for the strip", IEEE Trans. Ant. Propag., vol. 19, pp. 557-558, 1971.
9. G. D. Durgin, "Practical geometrical behavior of knife-edge diffraction", Proc. IEEE Ant. Propag. Soc. Int. Symp., San Diego, CA, July 5-11, pp. 1-4, 2008.
10. F. Hacivelioglu, L. Sevgi, P. Y. Ufimtsev, "Electromagnetic wave scattering from a wedge with perfectly reflecting boundaries: analysis of asymptotic techniques", IEEE Ant. Propag. Mag., vol. 53, pp. 232-253, 2011.
11. E. A. Essex, "Hertz vector potentials of electromagnetic theory", Am. J. Phys., vol. 45, pp. 1099-1101, 1977.
12. C. T. Tai, "Dyadic green's functions in electromagnetic theory", IEEE Press, New York, 1993.
13. M. Faryad, A. Lakhtakia, "Infinite-space dyadic green functions in electromagnetism", Morgan & Claypool Publishers, California, 2018.
14. S. W. Lee, "Comparison of uniform asymptotic theory and Ufimtsev's theory of electromagnetic edge diffraction", IEEE Trans. Ant. Propag., vol. 25, pp. 162-170, 1977.

15. B. L. Van der Waerden, "On the method of saddle points", *Appl. Sci. Res.*, vol. B2, pp. 33-45, 1952.
16. W. Pauli, "On asymptotic series for functions in the theory of diffraction of light", *Phys. Rev.*, vol. 54, pp. 924-931, 1938.
17. P. C. Clemmow, "Some extensions to the method of integration by steepest descents", *Quart. J. Mech. Appl. Math.*, vol. 3, pp. 241-256, 1950.
18. Y. Rahmat-Samii, R. Mittra, "A spectral domain interpretation of high frequency diffraction phenomena", *IEEE Trans. Ant. Propag.*, vol. 25, pp. 676-687, 1977.
19. Y. Rahmat-Samii, R. Mittra, "Spectral analysis of high-frequency diffraction of an arbitrary incident field by a half plane-Comparison with four asymptotic techniques", *Radio Sci.*, vol.13, pp. 31-48, 1978.
20. J. Boersma, Y. Rahmat-Samii, "Comparison of two uniform theories of edge diffraction with the exact uniform asymptotic solution", *Radio Sci.*, vol. 15, pp. 1179-1194, 1980.

Numerical Methods

21. G. Herrmann, "Numerical computation of diffraction coefficients", *IEEE Trans. Ant. Propag.*, vol. 35, pp. 53-61, 1987.
22. A. F. Peterson, S. L. Ray, R. Mittra, "Computational methods for electromagnetics", *IEEE Press*, New York, 1998.
23. T. Rylander, P. Ingelström, A. Bondeson, "Computational electromagnetics, 2nd ed.", *Springer*, New York, 2013.
24. R. S. Gupta, "Elements of numerical analysis, 2nd ed.", *Cambridge University Press*, Cambridge, 2015.
25. Z. Z. Chen, M. M. Ney, "Method of moments: a general framework for frequency- and time-domain numerical methods", in 2007 workshop on CEM-TD, Perugia, Italy, Oct. 15-17, pp. 1-4, 2007.
26. G. Apaydin, L. Sevgi, "Method of moments (MoM) modelling of wave propagation inside a wedge waveguide", *Appl. Comp. Electromag. Soc. J.*, vol. 29, pp. 515-522, 2014.
27. G. Cakir, L. Sevgi, P. Y. Ufimtsev, "FDTD modeling of electromagnetic wave scattering from a wedge with perfectly reflecting boundaries: comparisons against analytical models and calibration", *IEEE Trans. Ant. Propag.*, vol. 60, pp. 3336-3342, 2012.
28. M. A. Uslu, G. Apaydin, L. Sevgi, "Double tip diffraction modelling: finite difference

- time domain vs. method of moments", IEEE Trans. Ant. Propag., vol. 62, pp. 6337-6343, 2014.
29. O. Ozgun, L. Sevgi, "Finite element modelling of fringe fields in wedge diffraction problems", IEEE Ant. Propag. Letts., vol. 16, pp. 369-372, 2017.
 30. O. Ozgun, L. Sevgi, "Double-tip diffraction modelling: 2-D numerical models versus high-frequency asymptotics", IEEE Trans. Ant. Propag., vol. 63, pp. 2686-2693, 2015.
 31. F. Hacivelioglu, L. Sevgi, P. Y. Ufimtsev, "Wedge diffracted waves excited by a line source: exact and asymptotic forms of fringe waves", IEEE Trans. Ant. Propag., vol. 61, pp. 4705-4712, 2013.
 32. G. Apaydin, F. Hacivelioglu, L. Sevgi, P. Y. Ufimtsev, "Wedge diffracted waves excited by a line source: method of moments (MoM) modelling of fringe waves", IEEE Trans. Ant. Propag., vol. 62, pp. 4368-4371, 2014.
 33. G. Carluccio, F. Puggelli, M. Albani, "Generalization of UTD double diffraction to the case of impenetrable wedges with relatively general boundary conditions", IEEE Trans. Ant. Propag., vol. 62, pp. 3829-3834, 2014.
 34. G. Apaydin, L. Sevgi, "A novel wedge diffraction modelling using method of moments (MoM)", Appl. Comp. Electromag. Soc. J., vol. 30, pp. 1053-1058, 2015.
 35. M. A. Uslu, L. Sevgi, "MATLAB-based virtual wedge scattering tool for the comparison of high frequency asymptotics and FDTD method", Appl. Comp. Electromag. Soc. J., vol. 27, pp. 697-705, 2012.
 36. G. Apaydin, L. Sevgi, "Method of moments (MoM) modelling of wave propagation inside a wedge waveguide", Appl. Comp. Electromag. Soc. J., vol. 29, pp. 515-522, 2014.

FEM

37. O. C. Zienkiewicz, K. Morgan, "Finite elements and approximation", Wiley, New York, 1983.
38. O. C. Zienkiewicz, R. L. Taylor, J. Z. Zhu, "The finite element method: its basis and fundamentals, 6th ed.", Butterworth-Heinemann, Oxford, 2005.
39. S. C. Brenner, L. R. Scott, "The mathematical theory of finite element methods", Springer, New York, 2008.
40. M. G. Larson, F. Bengzon, "The finite element method: theory, implementation, and applications", Springer, Berlin, 2013.
41. P. Silvester, G. Pelosi, "Finite elements for wave electromagnetics: methods and

- techniques”, IEEE Press, New York, 1994.
42. J. P. Webb, “Application of the finite-element method to electromagnetic and electrical topics”, Rep. Prog. Phys., vol. 58, pp. 1673-1712, 1995.
 43. G. L. Wojcik, D. K. Vaughan, L. K. Galbraith, “Calculation of light scatter from structures on silicon surfaces”, Proc. SPIE, vol. 774, pp. 21-32, 1987.
 44. G. L. Wojcik, J. Mould *et al.*, “Numerical simulation of thick line width measurements by reflected light”, Proc. SPIE, vol. 1464, pp. 187-203, 1991.
 45. R. Capoglu, J. D. Rogers, A. Taflove, V. Backman, “The microscope in a computer: image synthesis from three-dimensional full-vector solutions of Maxwell’s equations at the nanometer scale”, Prog. Opt., vol. 57, pp. 1-91, 2012.
 46. <https://www.comsol.com>

5. Diffraction from Periodic Structures

This chapter extends modelling to include image formation in a microscope and explores the important relation between feature layout and the corresponding image. An earlier field simulation approach, the modal diffraction method, is first reviewed, together with methods for modelling image formation in an elementary microscope. Important field characteristics and the bright and dark field images of a single mesa on a silicon substrate simulated using the modal and the finite element method are then compared. Interference caused by the retro-reflected field from the substrate is found to be significant, so modifications are made to the simulation model. Finally, the effects of changing dimensions in a single and double mesa configuration are presented.

5.1 Modal diffraction model

In the last chapter, the numerical solution of the diffraction from an isolated mesa was presented, together with a chosen analytical model (GTD). However, the analytical model describes corner and half-plane diffraction only and provides no information about the effect of the dimension of the structure on the diffracted field. To provide this, Maxwell's equations must be solved for the diffraction problem of an isolated mesa. However, no apparent solution can be achieved for an isolated mesa as the problem cannot be accurately described. To obtain a solution, the structure is therefore commonly modelled as a periodic relief grating. Figure 5.1 shows simulation models of a) an isolated mesa and b) a mesa grating formed by repeating this structure with a periodicity of Λ . This modification allows the relative permittivity of the layer containing the mesa to be expressed as Fourier series, enabling a numerical solution of Maxwell's equations as the amplitudes of a set of diffraction orders.

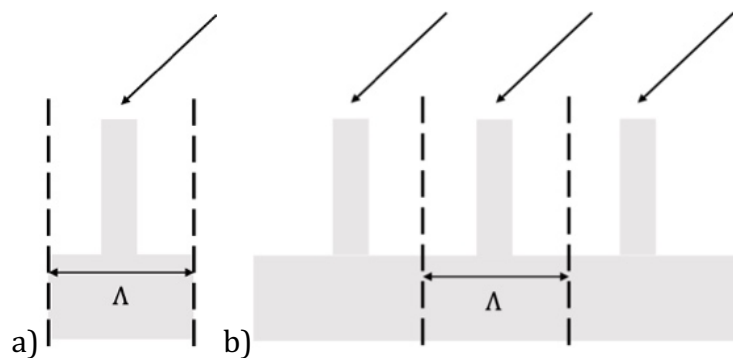


Figure 5.1: Simulation model of a) an isolated mesa and b) a periodic grating.

Periodic models have often been used to model object scattering in simulations of imaging. Examples include approximate diffraction theory [1, 2], modal theory [3-6] and rigorous coupled wave analysis [7-10]. The first is extremely inaccurate and cannot model structures that give rise to any significant in-plane scattering. The last two were originally developed to model volume holograms, and later adapted for surface relief gratings. Both are extremely accurate and might be expected to give a reasonable approximation of fields near isolated mesas, provided Λ is large enough that adjacent mesas do not scatter significantly onto each other. However, as with any numerical method, difficulties with convergence are encountered for large Λ . Both methods were originally developed to provide solutions for 2D problems, and later adapted to 3D. However, the computing resources needed in the latter case can be very high.

Once the scattered field is available, it can be used in a simple model of optical microscopy, with the effect of the lens being modelled simply as a modification of the diffraction order amplitudes and their recombination into an image. For example, Nyysönen considered optical imaging of photoresist lines in high-resolution optical lithography [11, 12] while Sheridan modelled edge detecting microscopy [13-15]. In each case the surface relief features varied only with a single coordinate, so the images were line images. Illumination was typically also monochromatic. Despite this simplification, computation times were long. Alternative methods were subsequently considered for plane images and white light illumination [16, 17], but the feature shapes involved were less appropriate for NEMS. Much later, Syms *et al.* studied optical imaging of HAR NEMS in bright and dark field using modal diffraction theory [18, 19], and it is this basic method that we follow here.

5.1.1 Single mesa

Figure 5.2a shows the simulation model used by Syms, which was based on a three-layer periodic relief structure consisting of 1) air; 2) a silicon mesa surrounded by air; and 3) a silicon substrate. Note that the coordinate system used was (z, x) , instead of the (x, y) used in the FEM model. The width w and depth d of the mesa are $0.1\mu\text{m}$ and $4\mu\text{m}$ respectively. The feature spacing Λ is $10\mu\text{m}$.

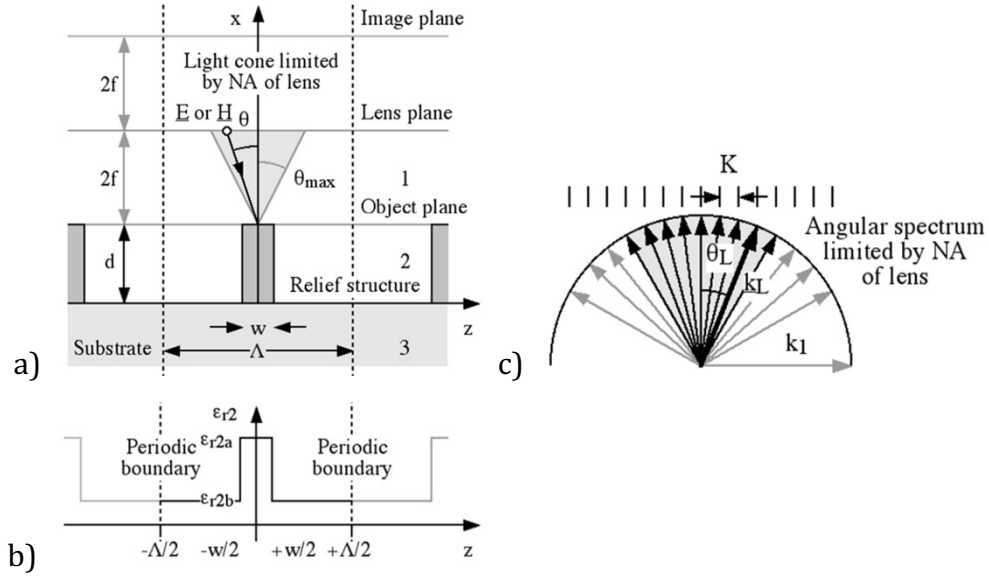


Figure 5.2: Simplified model of optical imaging [18]: a) overall system, b) assumed permittivity variation and c) effect of the lens on the spectrum of diffracted waves.

Figure 5.2b shows a section of the permittivity variation between $-\Lambda/2$ and $+\Lambda/2$. The relative dielectric constant of air ϵ_{air} is 1. For the silicon substrate and the mesa, the relative permittivity ϵ_{Si} was taken as $11.7 - 0.272j$ at $\lambda = 0.5\mu m$ [20, 21]. White light illumination was simulated by assuming plane wave incidence at a discrete number N_λ of wavelengths λ and discrete number N_θ of incident angles θ , with weighting by the aplanatic factor $\sqrt{\cos \theta}$. Bright and dark field illumination were both simulated, by appropriate choice of the angular range of illumination; Figure 5.2a shows bright field.

Each incident wave generates a back-scattered field, which consists of a fan of propagating and evanescent diffraction orders. To control the calculation, the maximum diffraction order was deliberately restricted (typically to $L_{max} = 250$). Figure 5.2c shows the angular spectrum of diffraction orders, which propagate at the corresponding diffraction angles θ_L . The grey region is the acceptance cone of the lens, which indicates the diffraction orders that can contribute to image formation. The larger the half-angle, the more diffraction orders can be collected and the higher the resolution of the image. Only propagating orders lying within the lens half-angle θ_{max} are collected. Here, the half-angle was assumed to be 30° .

In layer 2, the relative permittivity distribution ε_{r2} can be expressed with ε_{Si} and ε_{air} (referred as ε_{r2a} and ε_{r2b} in Figure 5.2b) as Fourier series with coefficients ε_{r2I} , so that:

$$\varepsilon_{r2}(z) = \sum_I \sigma_I \varepsilon_{r2I} \cos\left(\frac{I2\pi x}{\Lambda}\right) \quad (5.1)$$

$$\varepsilon_{r20} = \varepsilon_{Si} \frac{w}{\Lambda} + \varepsilon_{air} \frac{\Lambda - w}{\Lambda} \quad \varepsilon_{r2I} = \frac{\varepsilon_{Si} - \varepsilon_{air}}{I\pi} \sin \frac{I\pi w}{\Lambda} \quad (5.2)$$

Here σ_I is the Lanczos sigma function $\sigma_I = \text{sinc}\left(\frac{I\pi}{I_{max}+1}\right)$ [22]. This is used to reduce the ringing effect or Gibb's phenomenon [23] caused by the abrupt change of the relative permittivity ε_{r2} in layer 2 at $\pm w/2$. Since the permittivity variation is an even function, the Fourier series consists of cosine terms. The coefficients ε_{r2I} represents the I^{th} term in the Fourier series, also assumed to range from $-I_{max}$ to $+I_{max}$. When $I = 0$, the coefficient ε_{r20} is the DC term and expressed differently to other coefficients. Figure 5.3 shows the permittivity profiles of a single mesa illustrating the effect of the Lanczos sigma function. The solid lines (in black and grey) represent the exact permittivity variation of a single mesa and the dashed lines (in red and blue) show the permittivity profile computed in MATLAB. The right-hand figure shows that ringing effect is largely reduced at the feature, demonstrating the impact of the Lanczos function.

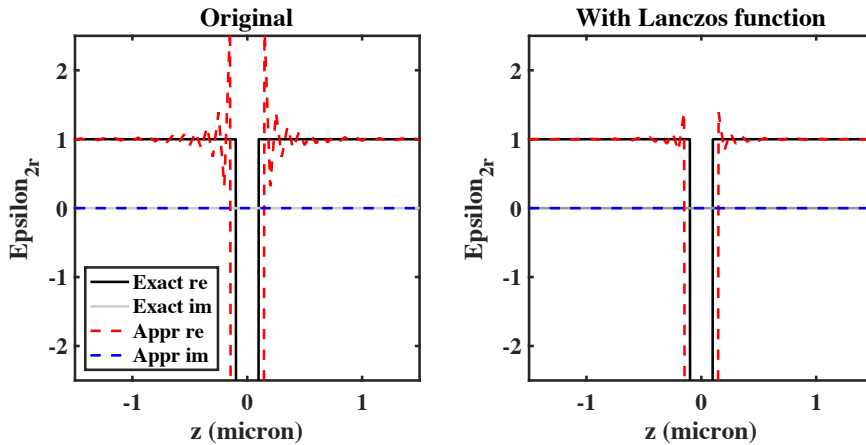


Figure 5.3: Permittivity profiles of a single mesa showing original expression (left) and improvement of Lanczos function (right).

For TE incidence, the wave equation is solved by assuming appropriate wave solutions and matching the tangential components of electric and magnetic fields at the boundaries between layers 1 and 2 and layers 2 and 3. Typically, this results in a large matrix problem. Once this has been solved, the solutions are combined in each layer as follows:

$$E_{y1} = e^{+jk_{x1}(x-d)}e^{-jk_z z} + \sum_L r_L e^{-jk_{x1L}(x-d)}e^{-jk_{zL} z} \quad (5.3)$$

$$E_{y2} = \sum_v [(A_v e^{+jk_{x2v} x} + B_v e^{-jk_{x2v} x}) \sum_L a_{Lv} e^{-jk_{zL} z}] \quad (5.4)$$

$$E_{y3} = \sum_L t_L e^{+jk_{x3L}(x-d)}e^{-jk_{zL} z} \quad (5.5)$$

The electric fields $E_{y1,2,3}$ represents the total field in each layer. In Equation 5.3, the first term is the incident wave. The remaining term is the back-scattered field and is the sum of a set of diffraction orders from $-L_{max}$ to $+L_{max}$ with amplitudes r_L . The propagation constant of the incident wave is k_0 , so that the z - and x - components are $k_0 \sin \theta$ and $k_0 \cos \theta$ respectively. For each diffraction order L , the z - and x -component (k_{zL} and k_{xL}) of the propagation constant are $k_{zL} = k_z + LK$ and $k_{xL} = \sqrt{\epsilon_{air} k_0^2 - k_{zL}^2}$. If k_{xL} is imaginary ($k_{zL}^2 > \epsilon_{air} k_0^2$), the corresponding diffraction order is evanescent.

Equation 5.4 describes the total field in layer 2 as a set of modes generated by reflection at the two interfaces. Here v is the mode number and k_{x2v} is the propagation constant of the v^{th} mode. A_v and B_v are the amplitudes of the forward and backward propagating waves v . The mode shapes are defined by the terms $a_{Lv} e^{-jk_{zL} z}$ when they are gathered for a particular v value. Equation 5.5 gives the total field in layer 3, which consists of a set of transmitted diffraction orders with amplitude t_L . By matching the electric field and its derivative at the layer boundaries, four sets of simultaneous equations with the unknowns r_L , t_L , A_v and B_v are obtained and are solved by matrix inversion. A unity magnification image may then be simulated by summing the back-scattered waves that fall inside the acceptance cone of the lens. A white light image may be generated by the incoherent sum of the image fields at all wavelengths and angles of incidence.

5.1.2 Double mesa

Other approximations to isolated structures can also be modelled using modal diffraction theory, simply by using the relevant permittivity expansion. This aspect may be used to explore in-plane scattering between closely spaced features, but there are clearly limits to the extent of each period that may be filled before scattering between features in adjacent periods becomes significant. One example is a double mesa, and Figure 5.4 shows a unit cell of a structure containing two similar adjacent mesas with a separation distance m at a period Λ . The mesa width and depth are the same as in Figure 5.2.

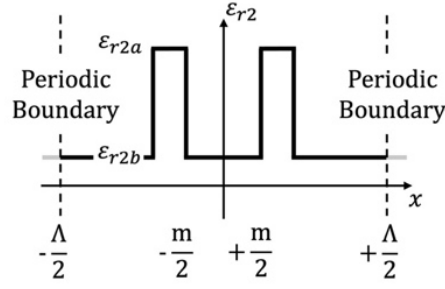


Figure 5.4: Permittivity variation of a double mesa model.

Since the structure is symmetric, the Fourier series of the permittivity profile is the same as the isolated mesa in Equation 5.1 but with different coefficients, that are expressed as:

$$\varepsilon_{r20} = \varepsilon_{r2Si} \frac{2w}{\Lambda} + \varepsilon_{r2air} \left(\frac{\Lambda - 2w}{\Lambda} \right) \quad (5.6)$$

$$\varepsilon_{r2l} = \left(\frac{\varepsilon_{r2Si} - \varepsilon_{r2air}}{l\pi} \right) \left[2 \cos \left(\frac{l\pi}{\Lambda} (m + w) \right) \sin \left(\frac{l\pi}{\Lambda} w \right) \right] \quad (5.7)$$

If m is zero, the coefficients reduce to the single mesa case with a feature width of $2w$ (Equation 5.2 and Figure 5.3).

5.2 Periodic finite element model

To set up a periodic model using the FEM, a rectangular box is created to define the total simulation region. Figure 5.5 shows a model of a silicon mesa with width w and depth d contained in a box of width Λ . The mesa sits on a silicon substrate and is otherwise surrounded by air. A perfectly matched layer (PML) is used at the top and bottom of the domain to absorb unwanted waves, enclosed by a “scattering boundary” condition. Since the evanescent field can be reflected by the boundary of the box, the domain height is chosen to allow sufficient decay of evanescent fields to avoid interference ($\sim 2\Lambda$). For the side boundaries, a Floquet boundary condition is used to account for periodicity.

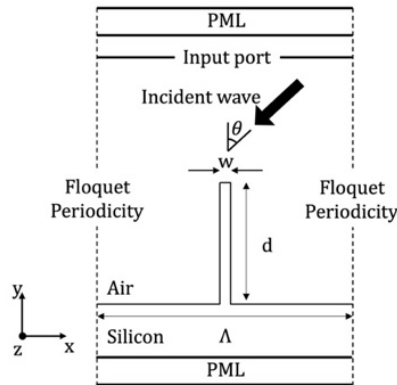


Figure 5.5: Simulation model for a periodic array of mesas in COMSOL.

An excitation port for the incident wave is assigned between the PML and the mesa. The black arrow shows the direction of the incident wave, which radiates towards the substrate with a power P_0 of 1W at an angle θ and wavelength λ . For most of what follows we assume $w = 0.1\mu m$, $d = 4\mu m$ and $\Lambda = 10\mu m$ as before, and that the relative permittivity of silicon is $\epsilon_{r_{Si}} = 11.7 - 0.272j$. We assume the incident angles of the illuminating waves are $0^\circ \leq \theta \leq 30^\circ$ for BF and $40^\circ \leq \theta \leq 60^\circ$ for DF. The half-angle of the lens is $\theta_{max} = 30^\circ$ in both modes.

5.2.1 Back-scattered field (E_s)

In COMSOL, the calculation for the total field is generated automatically, but the computation of the incident field is not straightforward. The total field above the mesa consists of the incident field and the back-scattered field (which contains both the reflected and diffracted field). The back-scattered field is a summation of the diffraction orders L of amplitudes r_L , ranging from $-L_{max}$ to $+L_{max}$. The full expression for a TE-polarised total field in z -direction is:

$$E_t = E_i + E_s = E_0 e^{+jk_y(y-d)} e^{-jk_x x} + \sum_L r_L e^{-jk_y L(y-d)} e^{-jk_{xL} x} \quad (5.8)$$

In the “Wave Optics” module of COMSOL, the incident port is excited at a defined power P_0 . In order to export the correct back-scattered field, the excitation power must be converted into electric field strength E_0 for the incident field. Generally, the energy w of an EM field is expressed as:

$$w = \frac{\epsilon_0 E^2}{2} + \frac{\mu_0 H^2}{2} = \epsilon_0 E^2 \quad (5.9)$$

Here E and H are related by the impedance (Equation 2.16) and as we are dealing with TE polarization, the energy is expressed by E field only. Let P be the energy transported per unit time across a unit cross-sectional area perpendicular to the direction of the wave. For a y -propagating plane EM wave, the wave travels a distance cdt in a time interval dt . Assuming this energy passes through a cross-sectional area A perpendicular to the y -axis, the EM wave sweeps through a volume of $dV = Acdt$. The energy in this volume is:

$$dw = w dV = \epsilon_0 E^2 Acdt \quad (5.10)$$

The power per unit area is then,

$$P = \frac{dw}{dt} \frac{1}{A} = \frac{\epsilon_0 E^2 Acdt}{Adt} = \epsilon_0 E^2 c \quad (5.11)$$

Since the average power per unit area transmitted by EM wave is $1/2$,

$$\bar{P} = \frac{\epsilon_0 E^2 c}{2} \quad (5.12)$$

Finally, the expression of the input electric (magnetic) field strength can be obtained as:

$$E = \sqrt{\frac{2\bar{P}}{\epsilon_0 c}} \quad H = \sqrt{\frac{2\bar{P}}{\mu_0 c}} \quad (5.13)$$

The full expression of the incident field in COMSOL is:

$$E_{inc} = \sqrt{\frac{2P_0}{\epsilon_0 c}} \frac{1}{\sqrt{A \times 1[m]}} e^{+j(k_0(y-y_0)\cos\theta - k_0 x \sin\theta)} \quad (5.14)$$

Here y_0 is the distance between the excitation port and the tip of the mesa. Note that P_0 is the total power excited at the incident port, and the average power per unit area is expressed as $\bar{P} = P_0/A$. The unit area A in a two-dimensional model (e.g. Figure 4.1) is the product of the width of the simulation model (Λ) and the depth of the simulation model in z -axis. The depth of the model is default 1[m] for a 2D model as the wave is propagating into the unit area on the x - z plane.

5.2.2 Amplitude of diffraction orders (r_L)

Now the back-scattered field can be calculated and exported from COMSOL. The data is then imported to MATLAB for calculating the amplitudes of the diffraction orders. From Equation 5.8, the back-scattered field is:

$$E_s = \sum_L r_L e^{-jk_{yL}(y-d)} e^{-jk_{xL}x} \quad (5.15)$$

Integrating both sides of Equation 5.15 with an exponential term $\exp(+jk_{xM}x)$ gives,

$$\int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} E_s e^{+jk_{xM}x} dx = \sum_L r_L e^{-jk_{yL}(y-d)} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{-j(k_{xL}-k_{xM})x} dx \quad (5.16)$$

For the integrand on the right-hand side, the exponential terms can be simplified as:

$$k_{xL} - k_{xM} = (k_x + LK) - (k_x + MK) = K(L - M) \quad (5.17)$$

Here K is the spatial frequency in the k -domain and is defined as $\frac{2\pi}{\Lambda}$. For $L - M \neq 0$, the integrand of the right-hand side of Equation 5.16 becomes:

$$\int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{-j(L-M)\frac{2\pi}{\Lambda}x} dx = -\frac{1}{j(L-M)\frac{2\pi}{\Lambda}} (e^{-j(L-M)\pi} - e^{+j(L-M)\pi}) = \frac{\Lambda}{(L-M)\pi} \sin((L-M)\pi) = 0 \quad (5.18)$$

For $L = M$, the integrand can be expressed as:

$$\int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} e^{-j(L-M)\frac{2\pi}{\Lambda}x} dx = \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} dx = \Lambda \quad (5.19)$$

Substituting the results into Equation 5.16 then yields:

$$\int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} E_s e^{+jk_{xL}x} dx = r_L e^{-jk_{yL}(y-d)} \Lambda \quad (5.20)$$

For focused images, the field must be calculated at the level of the tip of the mesa ($y = d$).

In this case, Equation 5.20 further reduces to:

$$\frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} E_s e^{+jk_{xL}x} dx = r_L \quad (5.21)$$

The amplitude of the diffraction order L can be obtained using Equation 5.21 by integrating the back-scattered field with the propagation term and dividing by the width of the simulation model. With this information, the image field can be reconstructed from the filtered diffraction orders.

5.2.3 Filtering of diffraction orders

To simulate the spatial filtering effect of the lens, the directions θ_L of each of the back-propagating diffraction orders must be known. These can be found by considering the x - and y -components of each propagation constant, namely:

$$k_{xL} = k_0 \sin \theta + LK \quad (5.22)$$

$$k_{yL} = \sqrt{k_0^2 - k_{xL}^2} = \sqrt{k_0^2 - (k_0 \sin \theta + LK)^2} \quad (5.23)$$

The diffraction angle θ_L is then given by,

$$\tan \theta_L = \frac{k_{xL}}{k_{yL}} = \frac{k_0 \sin \theta + LK}{\sqrt{k_0^2 - (k_0 \sin \theta + LK)^2}} \quad (5.24)$$

Since k_{yL} must be real for propagating orders, an inequality can be formed from Equation 5.24, as:

$$k_0^2 - (k_0 \sin \theta + LK)^2 \geq 0 \quad (5.25)$$

Expanding Equation 5.25 with the substitutions of k_0 and K , the inequality can be expressed as a quadratic equation in terms of L :

$$\frac{\cos^2 \theta}{\lambda^2} - \frac{2 \sin \theta}{\lambda \Lambda} L - \frac{1}{\Lambda^2} L^2 \geq 0 \quad (5.26)$$

The solutions give a range of diffraction orders L that satisfy the condition of real propagation constant k_{yL} at the incident angle θ . To simulate the effect of the lens by filtering out the diffraction orders outside its acceptance cone, the trigonometric relationship in Equation 5.24 can be used to compute the diffraction angles θ_L . The sine of the diffraction angle is therefore:

$$\sin \theta_L = \frac{k_{xL}}{k_0} = \frac{k_0 \sin \theta + LK}{k_0} = \sin \theta + \frac{L\lambda}{\Lambda} \quad (5.27)$$

The diffraction angle θ_L can be obtained by the arcsine of Equation 5.27. If we instead rearrange Equation 5.27 to make L to be the subject, the upper and lower limit of the accepted diffraction orders can be obtained with the substitution of $\theta_L = \pm 30^\circ$, that is the lens half-angle. Filtering of the diffraction orders can then be carried out by comparing each diffraction angle to the acceptance angle.

5.2.4 Image formation

The optical image of a structure can be calculated as the intensity of the image field, by reconstructing the image field from the filtered diffraction orders. However, an additional parameter must be considered, a correction factor that modifies the image based on the angle of incidence. When comparing images obtained at different angles, it is crucial to keep the incident power the same for all cases. Therefore, a factor of $\cos \theta$ is included in calculating the perpendicular input power, equivalent to a $\sqrt{\cos \theta}$ factor in the amplitude of the electric field. This is known as the illumination cosine factor. In COMSOL, the power of the excitation port is adjusted according to the angle of incidence.

A further factor, the aplanatic factor, is required on the image side. Figure 5.6a shows an imaging system consisting of an object space and the focal sphere. An incident ray at a distance h from the central axis travels through the object space and enters the focal sphere at an angle θ_d . Consider the object space in a 3D polar space, all the rays that are h away from the centre form a ring incidence. Any change in h transforms the ring into an annulus bounded by the circles of radius h and δh . Figure 5.6b shows such annuli for the incidence and image side, with $\delta S_{0,1}$ represents the area of the annuli respectively.

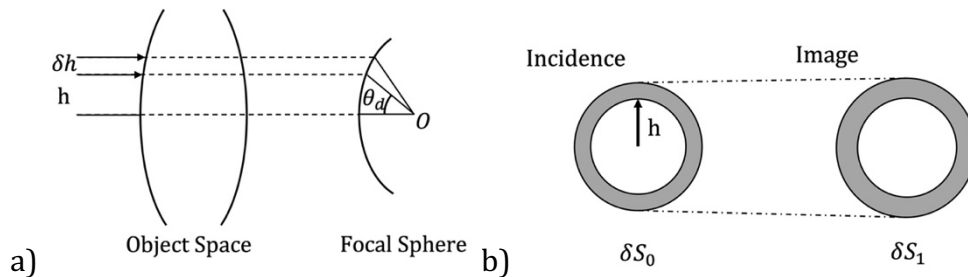


Figure 5.6: a) Meridional plane of an imaging system and b) annuli in incidence and image planes.

The electric fields in each space can be written as:

$$E_0 = l_0 \mathbf{e}_0 e^{-jk\zeta_0} \quad E_1 = l_1 \mathbf{e}_1 e^{-jk\zeta_1} \quad (5.28)$$

Here the subscript indicates the object (0) or the image (1) side. $\mathbf{e}_{0,1}$ is the unit vector in the direction of $E_{0,1}$ respectively, whereas $l_{0,1}$ are the amplitude factors and $\zeta_{0,1}$ are the corresponding values of the eikonal. Since the field is linearly polarised, the parameters are real-valued. According to the geometrical optics intensity law:

$$l_0^2 \delta S_0 = l_1^2 \delta S_1 \quad (5.29)$$

Here the refractive index of the object and image space is assumed unity. From Figure 5.6, the ratio of the areas δS_0 and δS_1 is $\cos \theta_d$. Therefore, Equation 5.29 can be rewritten as:

$$l_1 = l_0 \cos^{\frac{1}{2}} \theta \quad (5.30)$$

It means that the amplitude of the image field equals to $\cos^{\frac{1}{2}} \theta$ of the object field. Assume an object field is expressed as $E_{obj} = E_0 e^{-j(k_x x + k_y y + k_z z)}$, the image field obtained is [24]:

$$E_{imag} = E_0 e^{-j(k_x x + k_y y + k_z z)} \cos^{\frac{1}{2}} \theta_d \quad (5.31)$$

The optical image is then the incoherent sum of the square of the image field at different wavelengths and angles of incidence.

5.3 Numerical results

In the following section, numerical results obtained from COMSOL are compared with the predictions of the modal diffraction theory coded in MATLAB. First, diffraction from a single mesa with Si substrate is discussed, and bright and dark field images are shown for narrow and wide mesas. Modifications are then made to the simulation model so the diffracted field can be reviewed without the interference from the unwanted field components. The effects of altering the geometry of the feature for both single and double mesa configuration are discussed. TM polarization is not discussed as the magnetic field analysis is transmutable from the electric field analysis in TE polarization.

To avoid repetition, we focus on the bright field illumination for now. Dark field image is generated in a similar way, except for the wider range of incident angles. Figure 5.7 shows the electric field patterns of the diffraction from a single mesa of width $w = 0.1 \mu m$, depth $d = 4 \mu m$ and period $\Lambda = 10 \mu m$. Figures 5.7a and c show the result from the modal diffraction model and Figures 5.7b and d show the FEM results. The wavelength is $0.5 \mu m$. The incident angles are at 0° for the top row and at 20° for the bottom row, as indicated

by the red arrows. At normal incidence, the fringe patterns are symmetric and the patterns are a result of interference between the incident, reflected and diffracted waves. At oblique incidence, shadowing appears and the fringe pattern becomes asymmetric. Apart from some trivial differences, the results of different simulation methods are identical.

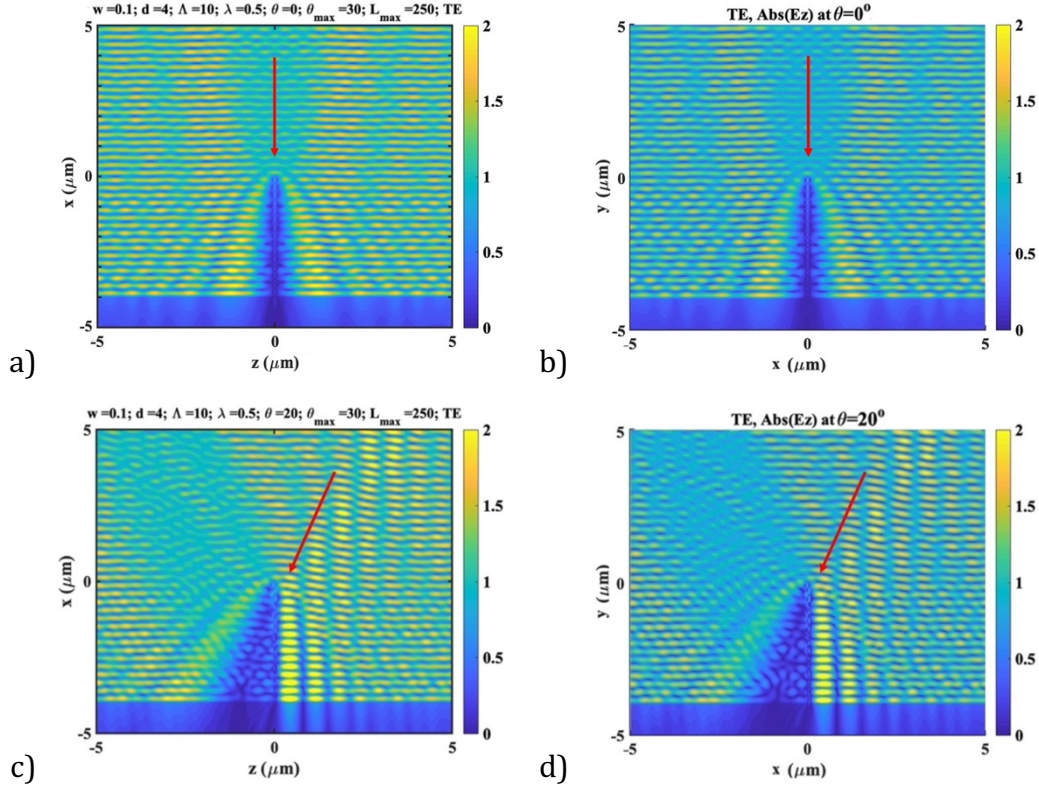


Figure 5.7: Total electric field for TE polarization at a) and b) 0° and c) and d) 20° incidence. The LH and RH columns shows results from modal diffraction and FEM.

Here we study the simulated results more closely by comparing the relative magnitude of the diffraction orders obtained from each simulation approach. We extract the electric field at the level of the tip of the mesa, the focus (imaging) level of the optical microscope. Diffraction order amplitudes are extracted from the field as discussed earlier. Figure 5.8 shows the relative magnitude of all considered diffraction orders at normal incidence by FEM and modal method in red and black lines respectively. Results are plotted on a log scale for the range of $-250 \leq L \leq +250$. In Figure 5.8a, the two sets of results show matching within the propagation zone, marked by dashed lines at $L_{p1} = -20$ and $L_{p2} = +20$. The result from the modal diffraction model is smooth and continuous. For FEM, the relative magnitude is approximately constant beyond $|L| = 100$ as the precision limit is reached. Figure 5.8b shows an expanded view of the diffraction order variation in

Figure 5.8a, focused on the range $-40 \leq L \leq +40$. The agreement of the relative magnitude of the diffraction orders is excellent for all propagating orders.

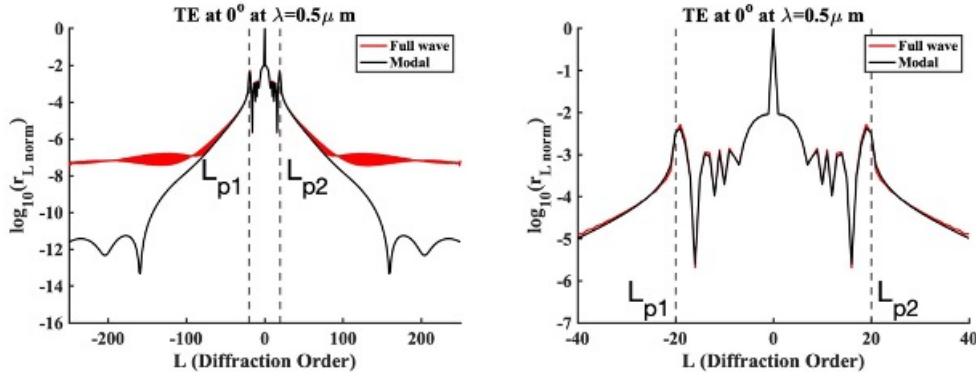


Figure 5.8: a) Overall relative amplitude variation of diffraction orders; b) Results for propagating zone.

We now plot the amplitude of the filtered diffraction orders in terms of their corresponding diffraction angles. Figure 5.9 shows the propagating diffraction orders against the diffraction angles from -90° to $+90^\circ$. The dashed lines θ_{f1} and θ_{f2} indicate the half-angle of the lens ($\theta_{max} = 30^\circ$). The magnitude of the diffraction orders simulated from both approaches are almost identical, as the black line lies on top of the red line for almost all diffraction angles. The strong peak at 0° is the specular reflection, which shifts to the side for oblique incidence. The diffraction orders within the acceptance range are collected for image reconstruction and those outside are filtered out. In general, the maximum practical lens half-angle is usually around 70° so the reconstructed image cannot fully resemble the object.

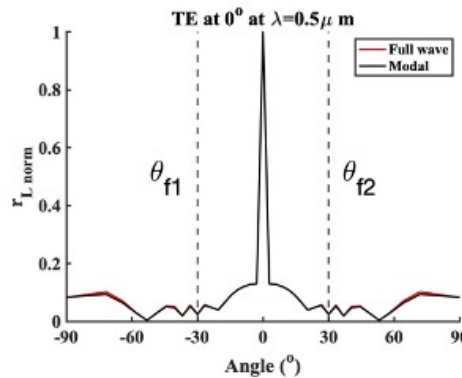


Figure 5.9: Angular spectra of relative diffraction order amplitude obtained using modal theory (black) and COMSOL (red).

Given the strong correspondence in the diffraction order spectra obtained using the two methods, the images predicted in each case should be similar. We now recombine the filtered diffraction orders to generate the image field. Figure 5.10 compares the object and image fields simulated by a) modal diffraction theory and b) full wave computation at the mesa tip. The object fields, black line in Figure 5.10a and red line in Figure 5.10b, show great similarity. The image (filtered) fields represented by the grey line in Figure 5.10a and blue line in Figure 5.10b are also analogous, particularly the large dip in the middle of the plot. Minor differences can be ascribed to the effect of the Lanczos filter used to smooth the permittivity variation in the modal model.

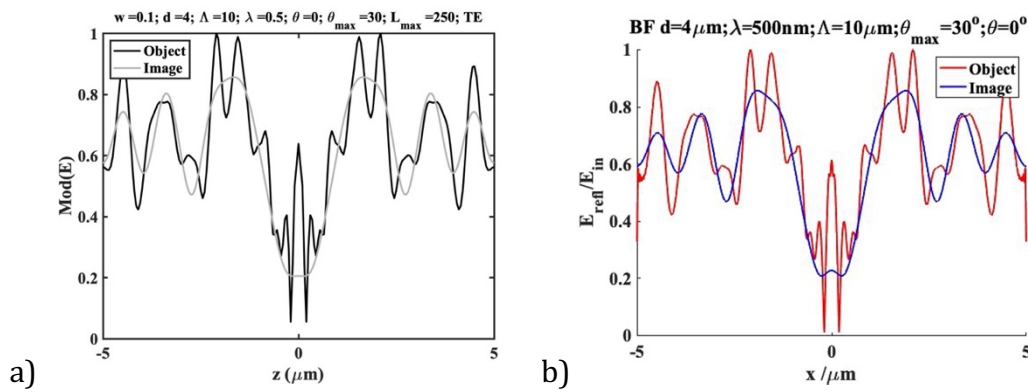


Figure 5.10: Object and image field at the imaging level for a) modal diffraction and b) FEM.

White light imaging can be simulated by summing the effect of incidence at different angles and wavelengths. Here we assume incoherent summation of the results obtained at 21 angles from 0° to 30° and 21 wavelengths between $0.4\mu m$ and $0.7\mu m$. Figure 5.11 shows the bright field images of mesas with feature width a) $w = 0.1\mu m$ and b) $w = 2\mu m$, representing NEMS and MEMS features, respectively. The black and red lines show the results obtained from the modal diffraction model and FEM respectively at different lens acceptance angles (dotted: 10° , dashed: 20° and solid: 30°). The grey models represent the mesas, and the dotted lines illustrate the level where the imaging takes place. For bright field imaging, the Si substrate appears as a bright background and features appear as dark lines. For both NEMS and MEMS features, there is a significant reduction in intensity at the feature edges, arising from the scattering of the incident waves into edge waves. As the acceptance angle increases, the angular range of these dips reduces, and the edge position are defined more accurately. In general, the intensity drop is larger for NEMS features, as the edge waves from the two side edges are augmented.

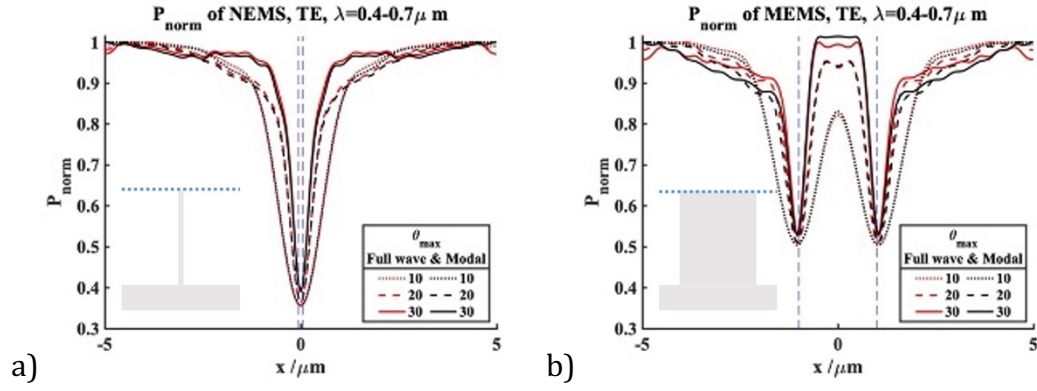


Figure 5.11: Bright field images of a) NEMS and b) MEMS features obtained from modal model and FEM, for different lens acceptance angles.

Similar analysis applies to dark field images, and the results agree with the bright field analysis. Figure 5.12 shows a) bright and b) dark field images of NEMS and MEMS mesas at a half-angle $\theta_{max} = 30^\circ$. The major difference between the BF and DF images is the background being bright ($P_{norm} = 1$) and dark ($P_{norm} = 0$) respectively. The feature itself then appears as a dip or a peak in the corresponding optical images. Figure 5.12a shows the difference in the intensity drop: for a NEMS feature, the dip is roughly 30% more than the MEMS feature. For dark field image, the MEMS peak is approximately half the NEMS peak. The contrast obtained in DF is much higher, allowing better feature visualisation. As a result, classification of NEMS and MEMS is easier in dark field [19]. However, oscillations are observed in both images, particularly for MEMS, as a retro-reflected field exists as a part of the back-scattered field.

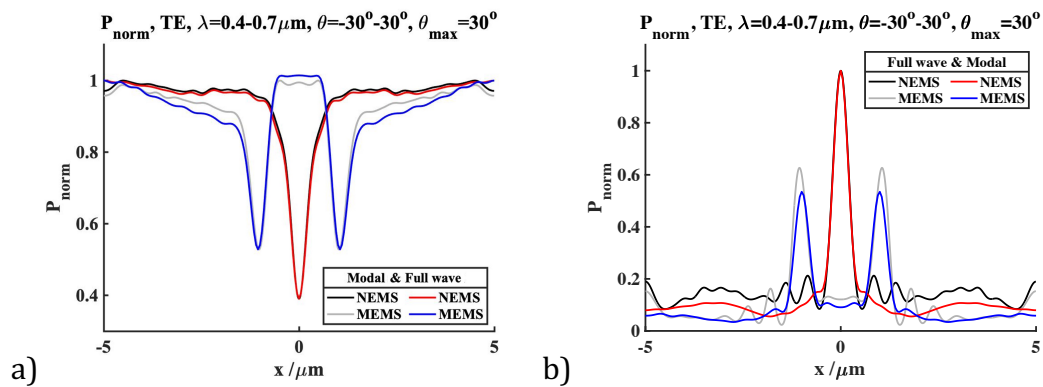


Figure 5.12: Bright and dark field images of mesas of feature width $0.1\mu m$ and $2\mu m$ generated from modal theory and FEM at $\theta_{max} = 30^\circ$.

5.4 Modifications

As seen above, many field components contribute to the white light image of a simple structure. The scattering problem is therefore complex, as the effects of each parameter cannot be isolated. However, the diffraction theories discussed in Chapter 4 suggest that the diffraction from a wedge corner is simply a cylindrical wave. Therefore, the simulation model was modified to simplify scattering problems in white light imaging.

5.4.1 Modified model

The aim of modifying the model is to eliminate unwanted field components so that the effect of altering the mesa dimension can be analysed. Three modifications were introduced: 1) illumination mode; 2) substrate and 3) feature material. In BF imaging, the reflected fields are collected by the lens together with the diffracted field. As the field strength of the reflected field is stronger, variations in the diffracted field are less pronounced. In DF imaging, only the edge wave is collected, and any effects on diffraction can be visualised easily. Similarly, the retro-reflected field contributes heavily to the BF image. The simplest way to remove this field is by removing the substrate, so the incident field reflects on the side of the mesa and travels in the direction opposite to the lens.

Figure 5.13 shows the field components for the above modifications. The black arrows represent the incident wave, and the red and green arrows represent the reflected and retro-reflected fields respectively. The blue circular wave front is the diffracted field. Figure 5.13a illustrates the reflected field entering the objective lens for bright field illumination. The specular reflected field in dark field is however travelling outside the objective lens. Figure 5.13b shows how the retro-reflected field in the scattering of mesa with substrate becomes a downward travelling reflected field.

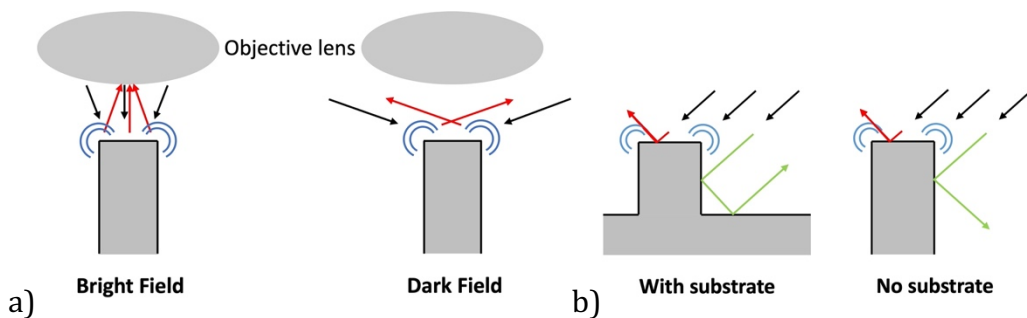


Figure 5.13: Field components between a) illumination modes and b) structures.

Lastly, since the waves propagate into silicon features, any field change is reduced by loss in the silicon. Therefore, replacing the feature with a perfect electric conductor (PEC) can ensure no field is lost during scattering. Figure 5.14 shows the modified simulation model of length Λ . PML boundaries are assigned on the top and bottom boundaries. The two sides are set as Floquet Periodicity boundaries as before. The feature material is by PEC with depth d and width w . The wave is incident from the top at angle θ , as indicated by the black arrow. We now examine results obtained using this modified model.

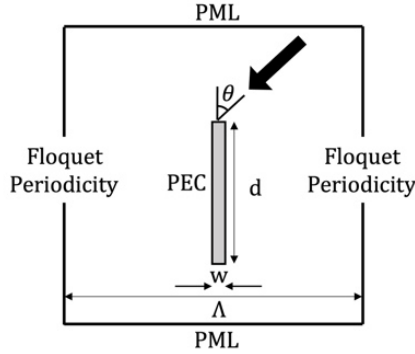


Figure 5.14: Modified FEM model.

5.4.2 Single mesa diffraction

Figure 5.15 shows the total field at a) normal incidence and b) $\theta = 20^\circ$ for the modified model. Figure 5.15a shows a generally weaker field strength compared to Figure 5.7a, especially in the region next to the feature. For oblique incidence, the standing wave pattern in Figure 5.7c on the illumination side of the feature is not seen in Figure 5.15b. The results confirm that the unwanted fields have been removed.

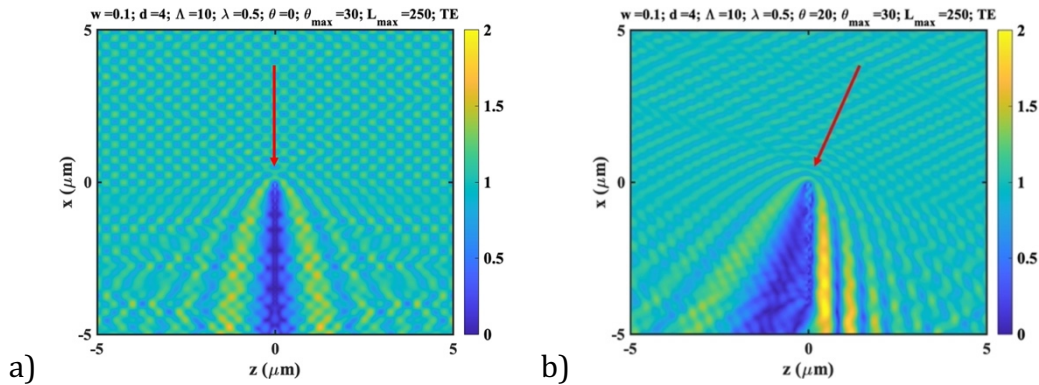


Figure 5.15: Total electric field for TE polarization at a) 0° and b) 20° incidence on a mesa with no substrate.

Figure 5.16 shows the object and image field of a single mesa in DF illumination at $\theta = 40^\circ$ along the imaging level, obtained from a) the modal model and b) FEM. The retro-reflected field is removed from the back-scattered field in the modified model, so the image field consists only the diffracted field (here, a two-dimensional equivalent of the Airy disc distribution). The results are substantially the same. Now the effect of the geometry of the feature can be addressed in terms of the change in magnitude or distribution of the image field.

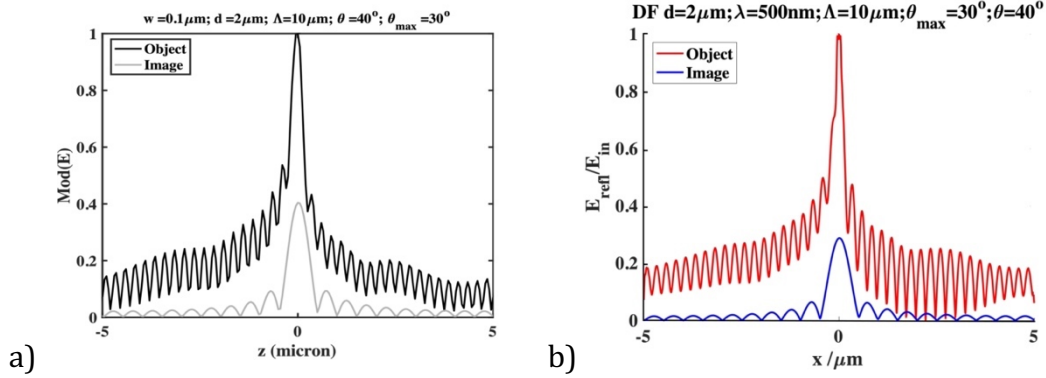


Figure 5.16: Object and image field variation simulated using a) modal diffraction and b) FEM for a mesa with no substrate for DF illumination at $\theta = 40^\circ$.

For the optical imaging of NEMS, it is important to understand the impact of changing the feature depth and width of the mesa. However, it is impractical to fabricate multiple NEMS devices for the analysis of the effects. Field simulations can be used instead. Since the field is obtained from a periodic model, neighbouring units can introduce unwanted interference. To investigate this effect, we varied the feature depth, feature width and the period as shows in Table 5.1. The default parameter set is $d = 2\mu m$, $w = 0.1\mu m$, $\Lambda = 20\mu m$ and $\lambda = 0.5\mu m$.

Parameter	Data points
Feature Depth ($d/\mu m$)	0.5, 2, 4, 8, 10, 16, 18, 20
Model Width ($\Lambda/\mu m$)	8, 10, 16, 20, 25, 30, 40, 50
Feature Width ($w/\mu m$)	0.001, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1

Table 5.1: Parameter list for the single mesa configuration.

Figure 5.17 shows the a) object fields and b) image fields of a single mesa with feature depths ranging from $d = 0.5\mu m$ to $20\mu m$, as stated in Table 5.1. The object fields for

mesas with different depths appear similar. After filtering by the lens, the image fields are virtually identical. These results imply that changes to the feature depth do not affect the image field, and consequently that the back-scattered field from the feature base does not give a significant contribution to the image.

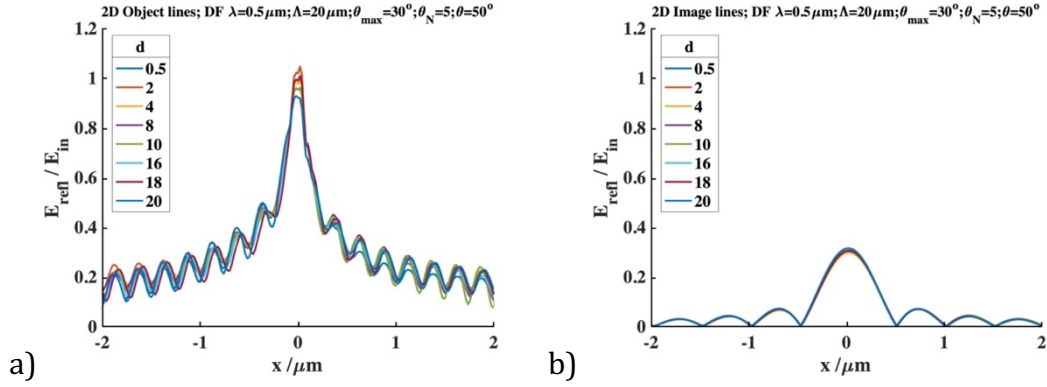


Figure 5.17: a) Object and b) image field of a single mesa at $\theta = 50^\circ$ for different feature depths.

Figure 5.18 shows similar comparison, but now varying the width of the simulation model over the range $\Lambda = 8\mu\text{m}$ to $50\mu\text{m}$. Oscillations in the object field away from the mesa varies significantly as the simulation model width increases, not probably due to a resonance between the neighbouring units when the period is a multiple of the horizontal equivalent wavelength ($\lambda_x = \lambda \sin \theta_0$). However, this resonance is filtered out by the lens as it only concerns horizontally propagating waves. As in the previous result, the image field is largely independent of the period of the simulation model.

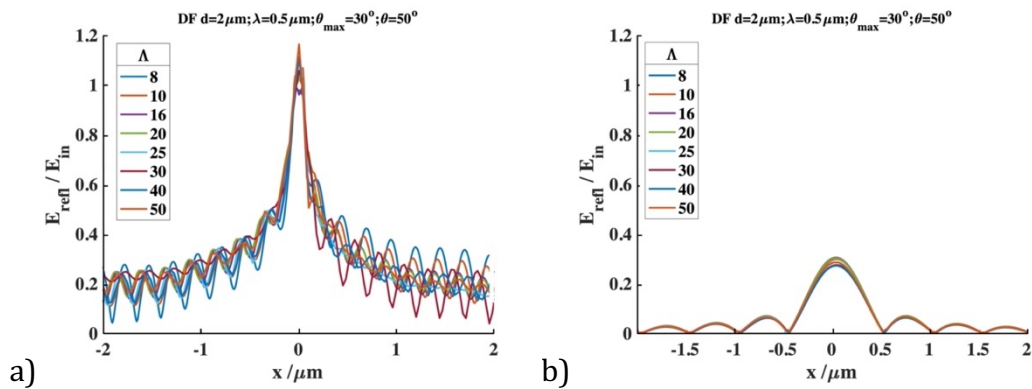


Figure 5.18: Electric field variation along the single mesa tip at $\theta = 50^\circ$ for a) object and b) image field, for different simulation model width.

Figure 5.19 shows the incoherent sum of the square of the image fields in dark field illumination at different wavelengths and angles of incidence for a) feature depth and b) model width. In Figure 5.19a, the intensity distributions from different feature depths are similar. The same observation can be made for Figure 5.19b as well, suggesting that the dark field image of a narrow mesa appears the same, regardless of the depth of the mesa or the size of the simulation space.

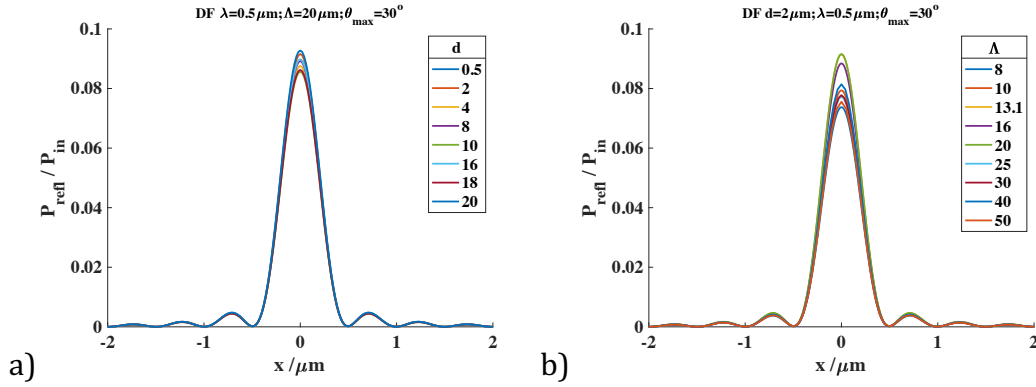


Figure 5.19: Light intensity variation for images of mesas with a) different feature depth d and b) different model width Λ .

Nonetheless, it is expected that the image field will vary with mesa width, following results in previous chapters. For narrow mesas, the edge waves generated from the two mesa corners add up in phase and appear as a single peak. For wider features, the edge waves appear as two distinct peaks. Figure 5.20 shows the light intensity variations for mesa widths ranging from $w = 0.001 \mu\text{m}$ to $1 \mu\text{m}$. For $w < 0.3 \mu\text{m}$, the intensity distributions are almost identical and the edge waves act as a single source. When w increases to $0.5 \mu\text{m}$, the central peak widens and the Abbe's diffraction limit is reached. Therefore, wider features are resolvable and are seen as two distinct peaks or two bright lines in DF images.

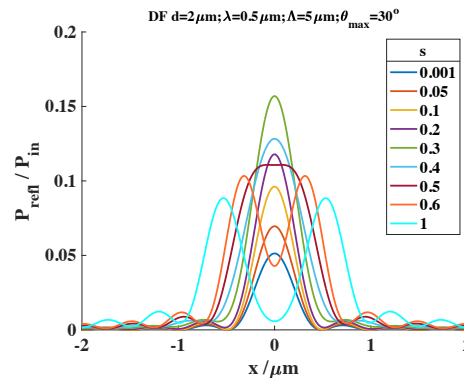


Figure 5.20: Light intensity variation for images of mesas with different width.

5.4.3 Double mesa diffraction

For a double mesa configuration, the modification of the simulation model in section 5.4.1 is also applicable. Figure 5.21 shows the total electric field of a double mesa, a) with and b) without a silicon substrate, at an incident angle $\theta = 50^\circ$. The separation between the mesas is $1\mu\text{m}$ and the mesa width is $0.1\mu\text{m}$. The standing wave pattern in Figure 5.21a is removed in Figure 5.21b and only the diffracted field remains.

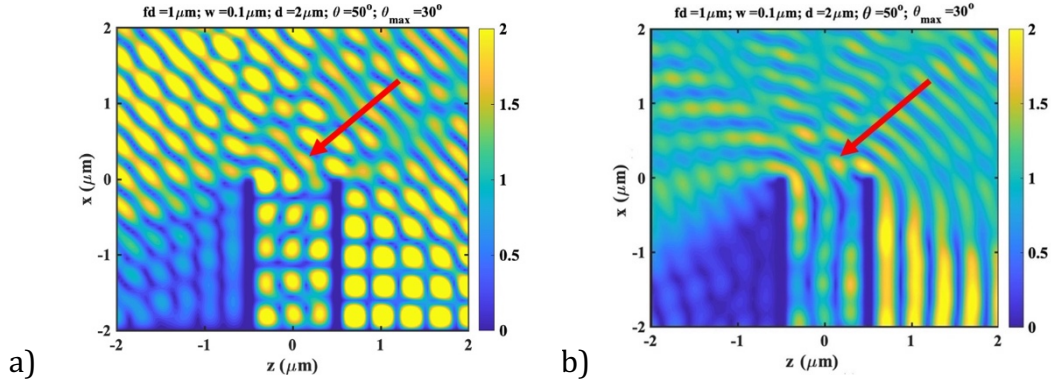


Figure 5.21: Comparison between a double mesa configuration a) with substrate and b) without substrate in an expanded view.

Figure 5.22 shows the object and image fields at the imaging level that are calculated from the modified model using a) modal diffraction and b) FEM. The object and image fields from both approaches are again very similar. The edge waves produced by each knife edge are seen as peaks in the plots, and radiate as cylindrical waves as expected.

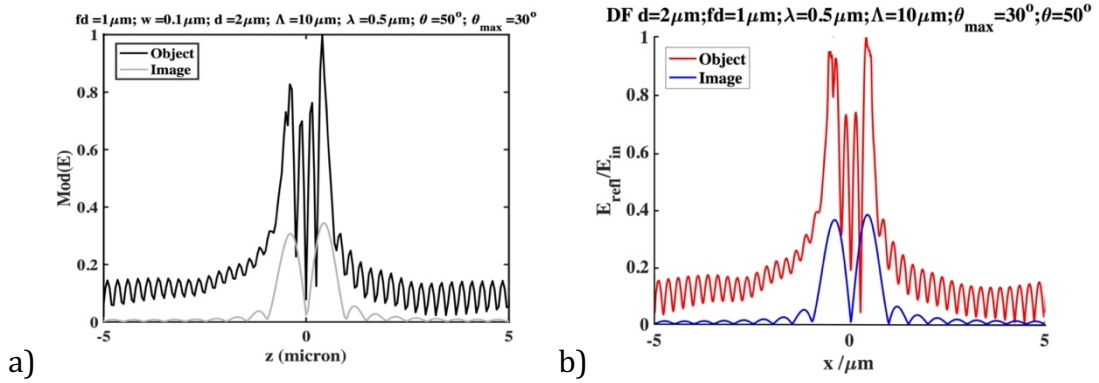


Figure 5.22: Object and image field calculated by a) modal model and b) FEM for a double mesa configuration with no substrate for DF illumination at $\theta = 50^\circ$.

The impact of changing the feature depth and width was also investigated for the double mesa configuration. Comparison of the results again shows that the image field is independent of feature depth. To avoid repetition, the results are omitted here. The effect

of varying the centre-to-centre separation s between the two mesas was also studied, using the parameters listed in Table 5.2.

Parameter	Data points
Separation Width ($s/\mu\text{m}$)	0.2, 0.3, 0.4, 0.5, 1, 1.5

Table 5.2: Parameter list for the double mesa configuration.

Figure 5.23 shows light intensity variations for different separations s ranging from $0.2\mu\text{m}$ to $1.5\mu\text{m}$ between two $0.1\mu\text{m}$ wide mesas. When s is small, the edge waves merge into a single peak. When s increases, separate intensity peaks are obtained but their apparent separation does not follow the spacing between the mesas. For example, the peak locations at $s = 0.4\mu\text{m}$ and $s = 0.5\mu\text{m}$ are very similar. Instead, amplitude varies. For example, the peak magnitude at $s = 0.5\mu\text{m}$ is increased by roughly 40%. The peak magnitudes also fluctuate significantly at small s . These effects can be ascribed to multiple diffraction between the mesas. However, multiple diffraction becomes negligible when $s > 1\mu\text{m}$, or when the mesas are at least two wavelengths apart. This effect can be attributed to weaker of the in-plane scattering, which falls off with distance.

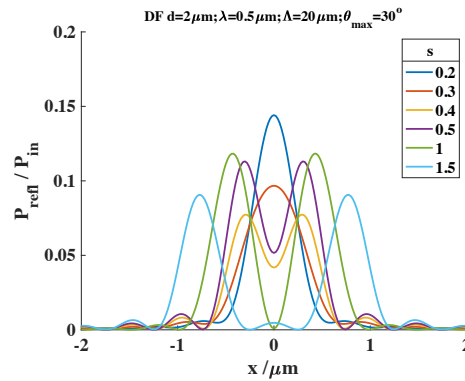


Figure 5.23: Light intensity for different centre-to-centre distance between a double mesa configuration.

Scattering from a double mesa is complex as it involves four edge waves and multiple diffraction. Now we examine the effect of the edge waves from the outer corners of the mesas by varying the mesa widths. Figure 5.24 shows the intensity of mesa pairs of width $w = 0.05\mu\text{m}$, $0.1\mu\text{m}$ and $0.2\mu\text{m}$ (red, green and blue lines respectively) at varying separations. Here the distance between the inner side edge of the double mesa is defined as m . The inner edges of the mesas are indicated as black dotted lines. Figure 5.24a-d shows the intensity for $m = 0.1\mu\text{m}$, $0.2\mu\text{m}$, $0.3\mu\text{m}$ and $1\mu\text{m}$ respectively. In each figure,

since the inner corners of the double mesa are at the same place, the difference in the intensity distribution must be caused entirely by the edge waves of the outer corners.

In Figure 5.24a, the intensity for $w = 0.05\mu\text{m}$ and $0.1\mu\text{m}$ appears as a single peak because the outer corners are close to the inner corners. Nonetheless, it is possible to resolve a double mesa with a tiny gap if the mesas are wider than $0.2\mu\text{m}$. At $m = 0.2\mu\text{m}$, double mesas are still unresolved at narrow feature. The double mesas with $0.3\mu\text{m}$ separation are then regarded as resolvable as two peaks are observed for all feature widths. Finally, if the separation is $1\mu\text{m}$ or more, the intensity distribution is formed solely by the edge waves and the effect of multiple diffraction diminishes.

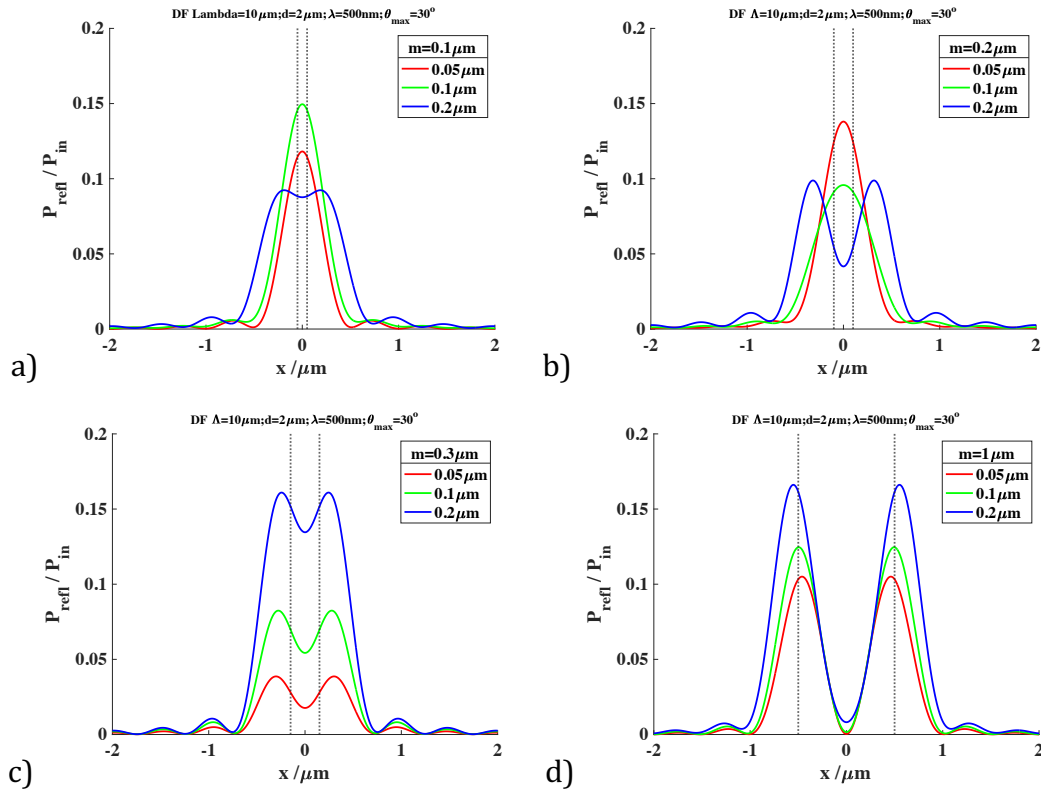


Figure 5.24: Light intensity variation for a double mesa configuration with different feature width and different mesa separation: a) $0.1\mu\text{m}$; b) $0.2\mu\text{m}$; c) $0.3\mu\text{m}$ and d) $1\mu\text{m}$.

5.5 Summary

Scattering from single and double mesas has been examined by comparing simulations obtained from FEM and modal diffraction theory. The results are in excellent agreement. Interference from the retro-reflected field has been addressed by removing the substrate from the model. The effect of changing feature depth and width has been examined by comparing field and intensity distributions in line images. For a single mesa, the

conclusion is that only a change in feature width can alter the apparent image. However, for a double mesa, an additional factor is the mesa separation. These findings highlight the most important parameters in optical imaging and allow simplification of the wave scattering problem. In the next chapter, 3D structures will be simulated in FEM and the optical images will be studied. This will establish the foundation for a novel imaging tool for NEMS devices.

References

Approximate models of object scattering

1. H. H. Hopkins, "On the diffraction theory of optical images", *Proc. R. Soc. A: Math. Phys. Sci.*, vol. 217, pp. 408-432, 1953.
2. E. C. Kintner, "Method for the calculation of partially coherent imagery", *Appl. Opt.*, vol. 17, pp. 2747-2753, 1978.

Modal theory

3. T. Tamir, H. C. Wang, A. A. Oliner, "Wave propagation in sinusoidally stratified dielectric media", *IEEE Trans. Microw. Theory Tech.*, vol. 12, pp. 323-335, 1964.
4. C. B. Burkhart, "Diffraction of a plane wave at a sinusoidally stratified dielectric grating", *J. Opt. Soc. Am.*, vol. 56, pp. 1502-1609, 1966.
5. F. G. Kaspar, "Computation of light transmitted by a thick grating, for application to contact printing", *J. Opt. Soc. Am.* Vol. 64, pp. 1623-1630, 1974.
6. R. S. Chu, J. A. Kong, "Modal theory of spatially periodic media", *IEEE Trans. Microw. Theory Tech.*, vol. 25, pp. 18-24, 1977.

Rigorous coupled wave analysis

7. M. G. Moharam, T. K. Gaylord, "Rigorous coupled-wave analysis of planar grating diffraction", *J. Opt. Soc. Am.* Vol. 71, pp. 811-818, 1981.
8. M. G. Moharam, T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings", *J. Opt. Soc. Am.* vol. 72, pp. 1385-1392, 1982.
9. M. G. Moharam, T. K. Gaylord, "Three-dimensional vector coupled wave analysis of planar-grating diffraction", *J. Opt. Soc. Am.*, vol. 73, pp. 1105-1112, 1983.
10. M. G. Moharam, T. K. Gaylord, "Rigorous coupled-wave analysis of grating diffraction - E-mode polarization and losses", *J. Opt. Soc. Am.*, vol. 73, pp. 451-455, 1983.

Simulation of 2D imaging

11. D. Niyssonen, "Theory of optical edge detection and imaging of thick layers", *J. Opt. Soc. Am.*, vol. 72, pp. 1425-1436, 1982.
12. D. Niyssonen, C. P. Kirk, "Optical microscope imaging of lines patterned in thick layers with variable edge geometry: theory", *J. Opt. Soc. Am. A*, vol. 5, pp. 1270-1280, 1988.
13. J. T. Sheridan, C. J. R. Sheppard, "Coherent imaging of periodic thick fine isolated

structures”, J. Opt. Soc. Am. A, vol. 10, pp. 614-632, 1993.

14. J. T. Sheridan, C. J. R. Sheppard, “Modelling of images of square-wave gratings and isolated edges using rigorous diffraction theory”, Opt. Comm., vol. 105, pp. 367-378, 1994.
15. J. T. Sheridan, T. O. Köurner, “Imaging periodic surface relief structures”, J. Microsc., vol. 177, pp.95-107, 1995.

Simulation of 3D imaging

16. S. van Haver, J. J. M. Braat *et al.*, “Vectorial aerial-image computations of three-dimensional objects based on the extended Nijboer-Zernike theory”, J. Opt. Soc. Am. A, vol. 26, pp. 1221-1234, 2009.
17. R. Capoglu, J. D. Rogers *et al.*, “The microscope in a computer: image synthesis from three-dimensional full-vector solutions of Maxwell’s equations at the nanometer scale”, Prog. Opt., vol. 57, pp. 1-91, 2012.

Imperial College imaging model

18. R. R. A. Syms, A. Bouchaala, O. Sydoruk, D. Liu, “Optical imaging and image analysis for high aspect ratio NEMS,” J. Micromech. Microeng., vol. 29, pp. 1-11, 2018.
19. R. R. A. Syms, A. Bouchaala, O. Sydoruk, “Improved optical imaging of high aspect ratio NEMS using dark-field microscopy”, Nanotechnology, vol. 30, pp. 1-8, 2019.

Silicon properties

20. K. Bücher, J. Bruns, H. G. Wagemann, “Absorption coefficient of silicon: an assessment of measurements and the simulation of temperature variation”, J. Appl. Phys., vol. 75, pp. 1127-1132, 1994.
21. M. A. Green, M. J. Keevers, “Optical properties of intrinsic silicon at 300K”, Prog. Photovolt., Res. Appl., vol. 3, pp. 189-192, 1995.

Lanczos function and the Gibbs phenomenon

22. C. Lanczos, “Discourse on Fourier series”, Edinburgh: Oliver & Boyd, Edinburgh, 1966.
23. D. Gottlieb, C. W. Shu, “On the Gibbs phenomenon and its resolution”, SIAM Review, vol. 39, pp. 644-668, 1997.

Aplanatic factor

24. B. Richards, E. Wolf, D. Gabor, “Electromagnetic diffraction in optical systems, II. Structure of the image field in an aplanatic system”, Proc. Roy. Soc. Lond. Math. Phys. Eng. Sci., vol. 253, pp. 358-379, 1959.

6. Simulation of 3D Images

This section presents models of optical microscope images of different 3D structures and studies the effect of illumination and topology. A 3D FE model is first developed from the 2D model of the previous chapter. Its predictions are then validated by comparison with the 2D FE and modal simulations for TE polarization. The effect of illumination angle is identified from examples of a high aspect ratio prisms and cylinders, and the results are used to construct a simpler model in the following Chapter. Simulated optical images of different canonical HAR structures are then presented, and interesting structures that produce highly similar images are discussed. Finally, the effect of additional incident angles and wavelengths is considered, together with TM polarization.

6.1 Scattering from 3D structures

Optical imaging of 3D microscale features has received considerable attention over the last few decades [1-4]. The object was first modelled as a transmission function [5-7] and different diffraction theories were developed [8]. Scattered fields were then computed more accurately by finite element (FE) [9-11] and finite difference time-domain (FDTD) [12, 13] methods. Other approaches have included rigorous coupled wave theory (RCWA) [14, 15], vector potentials [16, 17] and vectorial diffraction theory [18]. Canonical diffraction problems (for example, 3D half-planes) have also been studied [19, 20]. However, a common problem is the heavy computational resources needed. Simulation times are therefore always much larger than in 2D, and there is always an adverse trade-off between model size and computation time. For a 3D structure, the scattering problem is more complicated as the angular spectrum of diffraction orders comprises a cone instead of a fan. Due to the very large number of orders involved, attention here is restricted to a 3D FEM model, because of the difficulty of solving the increased number of simultaneous equations generated by boundary matching in the 3D modal method. Similarly, the object is assumed to be formed from PEC surfaces throughout, to avoid complications introduced by materials.

Figure 6.1a shows the field components involved in scattering from a 3D PEC mesa. Here the incident field (red arrow) illuminates the mesa at a vertical angle θ_{in} from the +y-axis and makes a horizontal (azimuth) angle ϕ_{in} with the -z-axis. The two common polarizations have E_{in} and H_{in} in parallel to the x-z plane and are denoted as TE and TM

polarizations here. In each polarization, the second field vector (H_{in} in TE or E_{in} in TM) is perpendicular to both the propagation direction and other field vector (blue arrows). The green arrow represents the reflected field with a reflection angle θ_{refl} from the $+y$ -axis at a horizontal angle ϕ_{in} to the $+z$ -axis. The orange arrow represents one of the diffracted rays, with a vertical angle θ_{diff} . Figure 6.1b shows the FEM model of a mesa in 3D simulation space. In the $x - y$ plane, the structure is constructed the same way as the 2D model. However, it is now extruded along the z -axis with a period of Λ_2 .

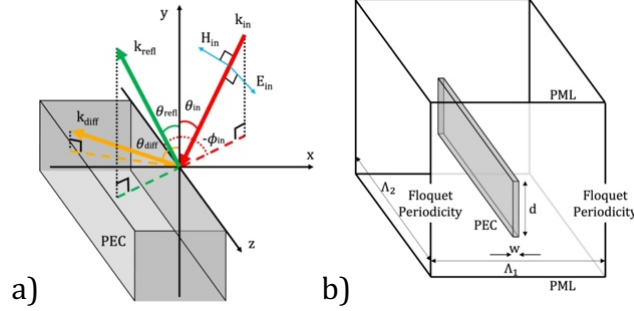


Figure 6.1: 3D Mesa models: a) wave components for TE polarization; b) 3D FEM model of a periodic mesa.

The x - and z -components of the incident electric (magnetic) field in TE (TM) polarization vary with the azimuth angle ϕ_{in} as shown in Equation 6.1 [21]:

$$\mathbf{E}_0 = \begin{bmatrix} \cos \phi_{in} \\ 0 \\ -\sin \phi_{in} \end{bmatrix} \quad (6.1)$$

In TE polarization, the total field E_t in the 2D scattering problem is expressed as [22-24]:

$$\mathbf{E}_t = \mathbf{E}_0 e^{-jk_{y0}(y-d)} e^{-jk_{x0}x} e^{-jk_{z0}z} + \sum_L \sum_P \mathbf{r}_{L,P} e^{-jk_{yL,P}(y-d)} e^{-jk_{xL,P}x} e^{-jk_{zL,P}z} \quad (6.2)$$

$$k_{x0} = k_0 \sin \theta_{in} \sin \phi_{in} \quad k_{z0} = k_0 \sin \theta_{in} \cos \phi_{in} \quad k_{y0} = -k_0 \cos \theta_{in} \quad (6.3)$$

Here the first term in Equation 6.2 is the incident field $\mathbf{E}_i$, and the second term is the back-scattered field $\mathbf{E}_s$ obtained by summing all diffraction orders with their corresponding amplitudes $\mathbf{r}_{L,P}$. The terms k_{x0} , k_{y0} and k_{z0} in Equation 6.3 are the x -, y - and z -components of the propagation constant of the incident field. Similarly, the terms $k_{xL,P}$ and $k_{zL,P}$ are the x - and z -components of the propagation constants of the diffraction orders [25, 26]. These components are defined as:

$$k_{xL,P} = k_{x0} + K_x L \quad k_{zL,P} = k_{z0} + K_z M \quad k_{yL,P} = \sqrt{k_{0,0}^2 - k_{xL,M}^2 - k_{zL,M}^2} \quad (6.4)$$

To limit the calculation, the indices L and P are assumed to range from -250 to +250. For TM polarization, the same approach is used to find the back-scattered amplitudes $\mathbf{r}_{L,P}$.

As before, the image field is generated by allowing only the diffraction orders lying inside the acceptance cone of the objective lens to contribute. These orders must be propagating, and all information in evanescent orders will therefore be lost. Figure 6.3a shows the acceptance cone (in dashed lines) superimposed on an example of a thin mesa. The red arrow indicates the incident wave, and the black arrows are the diffraction orders lying inside the acceptance cone. Figure 6.3b shows a 2D representation of the diffraction orders that lie inside the cone, assuming incidence at $\theta = 0^\circ$ and $\phi = -90^\circ$. In this view, the result of filtering is presented as a polar plot, retaining only the orders lying inside a circle. The diffraction order distribution is uniform because the increment (decrement) is constant in both x - and z -directions.

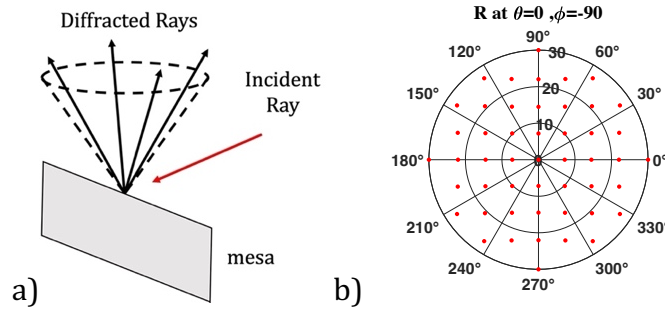


Figure 6.3: Effect of the lens: a) acceptance cone; b) diffraction orders within the cone.

Since the scatterings in the x - and z -directions are independent, the 2D problem can be reduced to two separate 1D problems. The total fields in the x - and z -directions are first extracted separately from COMSOL, and the corresponding incident fields are then subtracted. For TE polarization, the scattered field in each direction is expressed as:

$$x: E_s = E_t - E_i \cos \phi \quad z: E_s = E_t + E_i \sin \phi \quad (6.5)$$

For TM polarization, the electric field components are simply replaced by magnetic fields. Next, the x - and z -components of the amplitudes of the diffraction orders $\mathbf{r}_{L,P}$ are calculated by double integration along the x - and z -directions:

$$r_{x,z,L,P} = \frac{1}{\Lambda_1} \frac{1}{\Lambda_2} \int_{-\frac{\Lambda_1}{2}}^{\frac{\Lambda_1}{2}} \int_{-\frac{\Lambda_2}{2}}^{\frac{\Lambda_2}{2}} E_{x,z,S} e^{+j(k_{xL}x + k_{zP}z)} dx dz \quad (6.6)$$

Here $E_{x,z,S}$ represents the back-scattered field in the x - and z -directions, respectively. Separate images obtained from the x - and z - field components are then reconstructed by summing the filtered diffraction orders, for simplicity neglecting image inversion. Finally, the complete TE image is obtained from the sum of these two image intensities. A complete TM image is generated from the same process.

6.2 Comparison with 2D Model

We begin the 3D result analysis by validating the total electric field with the 2D models discussed in the previous chapter. For example, Figure 6.4 shows the total electric field scattered from a thin mesa of width $0.1\mu\text{m}$ and depth $2\mu\text{m}$ in a $6\mu\text{m}$ wide simulation space. This space has been considerably reduced compared to the FE models in the previous chapters to accommodate the larger computation demands of 3D. The first three sub-plots show 2D field variations of a thin mesa (as grey blocks) obtained using the 2D modal model (6.4a), 2D FEM (6.4b) and 3D FEM (6.4c) at 50° incidence, for TE polarization. In each case the field patterns are essentially identical. The last sub-plot (6.4d) compares line object and image fields at the mesa tip for the 2D modal and 3D FEM models. Agreement between the 2D modal model and the 3D FEM is excellent. These results confirm that 3D FE model is reliable, despite the reduction in model size.

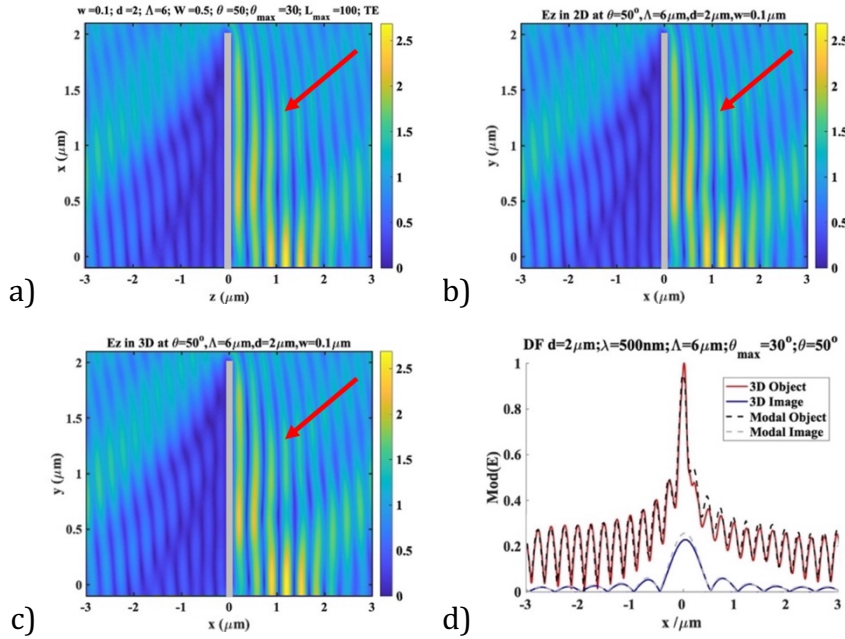


Figure 6.4: Simulated electric field of a single thin mesa at $\theta_{in} = 50^\circ$ (and $\phi_{in} = 90^\circ$) obtained using a) 2D modal model, b) 2D FEM and c) 3D FEM. d) Object and image fields obtained using 2D modal model and 3D FEM. TE polarization.

6.3 Simulation of 3D imaging

A major difference between 3D and 2D models is the addition of an azimuth angle for the incident wave. Figure 6.5a illustrates this angle for a thin HAR mesa. The red arrows represent the direction of the incident wave. When this is altered by rotating the wave vector about the y -axis, the strength of the back-scattered field varies significantly. In fact,

it is well known that scattering is maximised when illumination is perpendicular to the edge [27, 28]. Consequently, when structures such as the cylinder in Figure 6.5b are illuminated by a single wave, only parts that are perpendicular to the incident field generate significant scattering; others hardly contribute. However, scattering from the regions shown must involve the leading edge in one case and the trailing edge in the other.

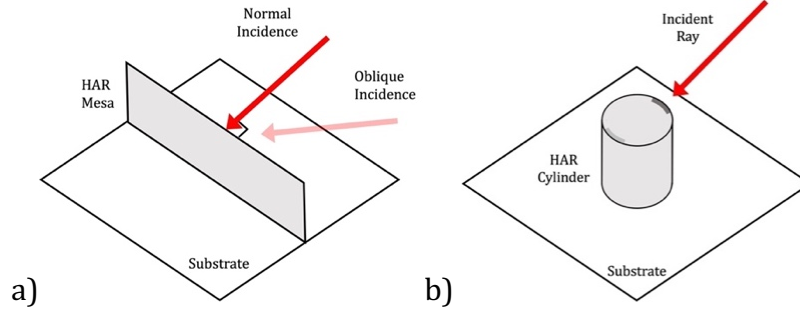


Figure 6.5: Illumination of a) HAR mesa and b) HAR cylinder.

6.3.1 Rectangular structures

We demonstrate the dominance of normal illumination by simulating the scattering from rectangular structures of a single wave incident perpendicular to an edge. Figure 6.6 shows the illumination of a) a solid and b) a hollow HAR square prisms. The highlighted edges show where most of the scattering takes place. In Figure 6.6a, the incident wave is scattered by just two edges of the prism. Scattered waves are generated at every point along these edges, and add to give cylindrical edge waves, except at the corners. However, scattering is more significant at the leading edge, generating a brighter image in DF. Scattering from the hollow square prism takes place in similar regions. However, since both regions contain leading and trailing edges, the image will be similar.

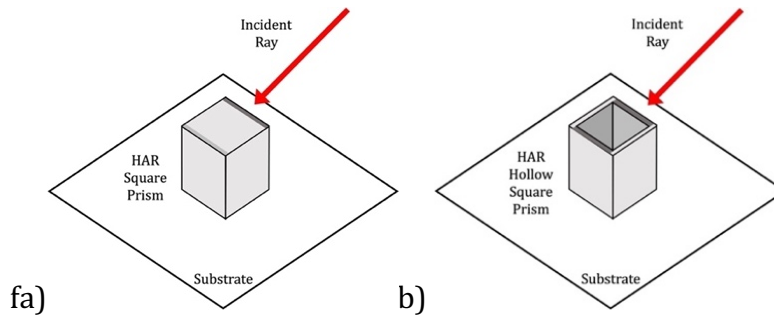


Figure 6.6: Illumination of a) solid and b) hollow HAR square prisms by a single wave.

Figure 6.7 shows TE DF FEM images of a solid and hollow square prisms of side length $5\mu\text{m}$ and depth $1\mu\text{m}$ in a $8\mu\text{m}(x) \times 8\mu\text{m}(z) \times 2\mu\text{m}(y)$ simulation space. The sidewall

thickness of the hollow prism is $0.1\mu\text{m}$. Figure 6.7a shows the image of a square prism with a single wave incident at $\theta_{in} = 50^\circ$ and $\phi_{in} = -90^\circ$ (red arrow). The leading edge of the structure is bright and the trailing edge is dim; the other two edges are barely visible. Figure 6.7b shows the corresponding image for a hollow square prism. Now both vertical edges appear bright, because both contain leading and trailing edges. The result of more general illumination can be simulated by incoherently summing the effects of additional incident waves. Figure 6.7c and d show the images of solid and hollow square prisms obtained using four incident waves ($\phi_{in} = -90^\circ, 0^\circ, 90^\circ, 180^\circ$). Both images now show a complete outline. However, the corners are rounded and dim.

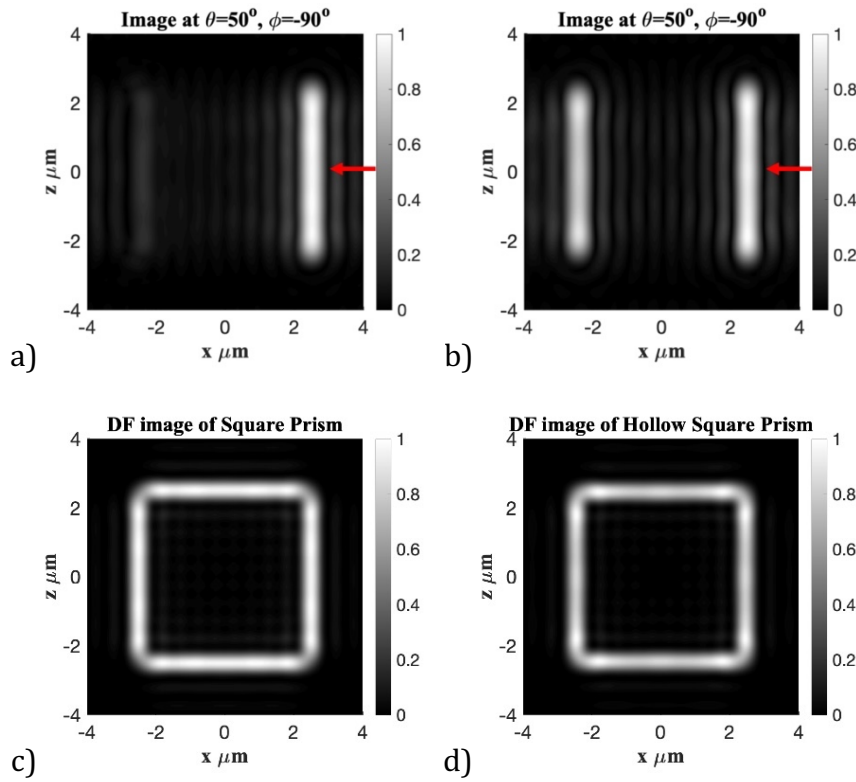


Figure 6.7: DF FEM Images of a) solid and b) hollow square prisms for $\theta = 50^\circ$ and $\phi_{in} = -90^\circ$; c) and d) corresponding images for 4 illuminating waves. TE polarization.

The absence of corners is a consequence of illumination by a small number of waves. However, Figure 6.8a shows the image of a solid square prism for a single wave incident at $\theta_{in} = 50^\circ, \phi_{in} = -45^\circ$. Now the corners are visible as bright circular spots, and the edges are missing. This result suggests that corners may be restored by including oblique waves. Figure 6.8b shows the image obtained using eight waves incident at angles from $\phi_{in} = -90^\circ$ to 225° in steps of 45° . We conclude that as few as four illuminating waves

can provide a reasonable result for Manhattan geometries, and that eight waves provide a marginal improvement at corners.

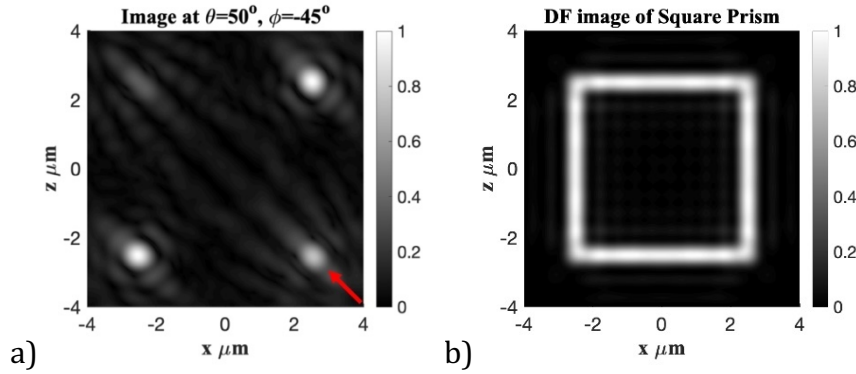


Figure 6.8: a) DF image of a square prism at $\theta_{in} = 50^\circ$ and $\phi_{in} = -45^\circ$ by FEM; b) corresponding image obtained using 8 illuminating waves. TE polarization.

6.3.2 Curved structures

For structures with curved surfaces, the image field must again be maximised in regions where the edge is perpendicular to the incident wave and reduce gradually as the angle between the wave and edge varies. Figure 6.9 shows the illumination of a) solid and b) hollow HAR cylinders. In each case, the scattered field must concentrate in the shaded regions where the edge is approximately perpendicular to the incident wave.

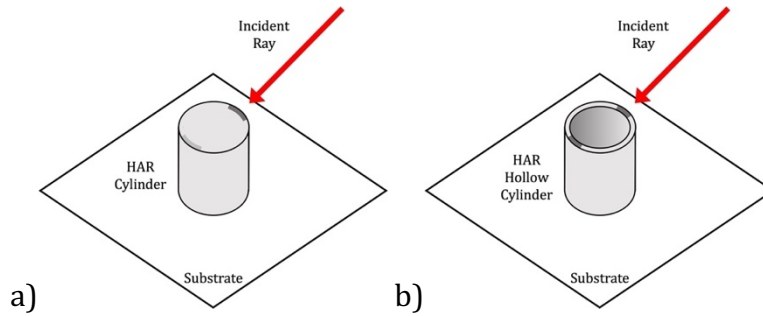


Figure 6.9: Illumination of a) solid and b) hollow HAR cylinders.

Figure 6.10 shows simulated TE images of solid and hollow cylinders of radius $2.5\mu\text{m}$ and depth $1\mu\text{m}$ which confirm this assumption. Figure 6.10a shows the image of a solid cylinder with a single wave incident at $\theta_{in} = 50^\circ$ and $\phi_{in} = -90^\circ$ (red arrow). The image consists of two arc segments corresponding to leading and trailing edge diffraction at near perpendicular incidence. In between, the structure is missing. Figure 6.10b shows the image of a hollow cylinder; now both arcs are of equal brightness, since both contain leading and trailing edges. Figures 6.10c and 6.10d show the corresponding images with

four waves incident. Now the structure is visible as four arcs, suggesting that complete ring images would be obtained using only a few more waves. Figures 6.11a and b, which show images obtained using eight waves, confirm this assumption. These results suggest that diffuse illumination of general structures can be modelled with a limited number of incident waves, and good results can be achieved with just eight [29-32].

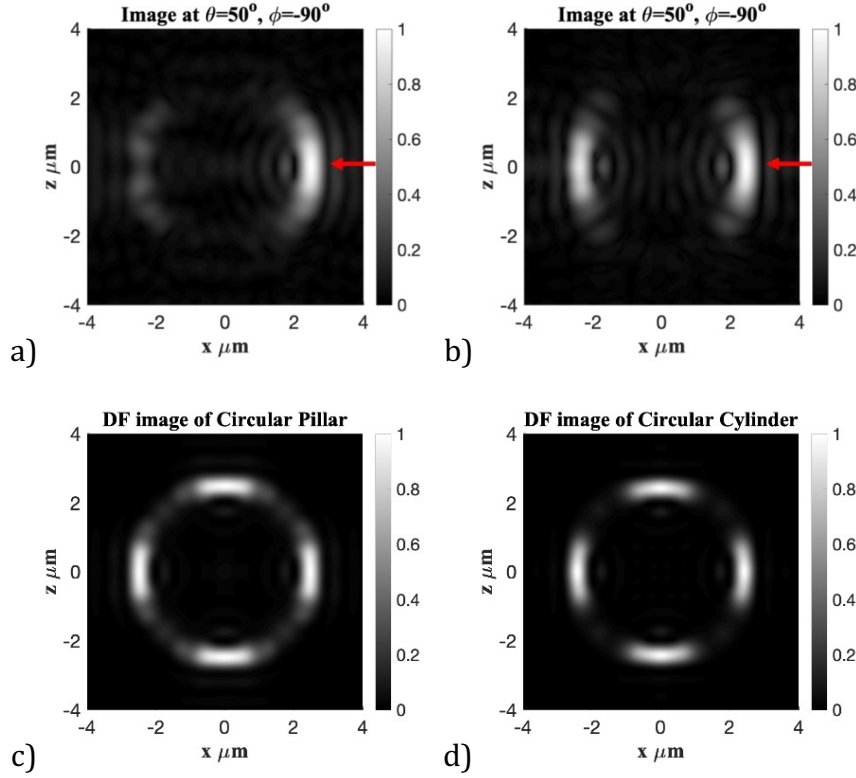


Figure 6.10: DF images of a) solid and b) hollow cylinders for incidence at $\theta_{in} = 50^\circ$ and $\phi_{in} = -90^\circ$; corresponding images using four illuminating waves. TE polarization.

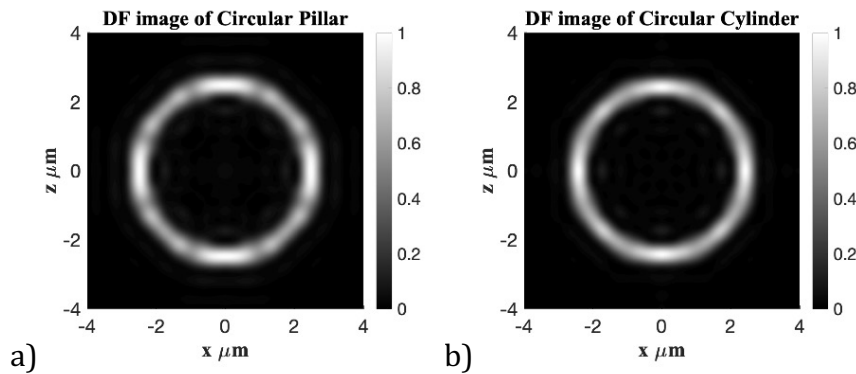


Figure 6.11: DF images of a) solid and b) hollow cylinders for eight waves. TE polarization.

6.4 Image library

Simulated DF images of canonical structures from TE polarization are now presented. Given the high quality of monochromatic images obtained using four waves incident at azimuth angles ϕ_{in} from -90° to $+180^\circ$, the same approach is used, and images are normalised to enhance contrast. To reduce the simulation time, all models are restricted to $1\mu\text{m}$ depth.

6.4.1 Single mesas

The first model is a single mesa of infinite length oriented along the z-axis. Figure 6.12 shows a) the layout and b) the image of a nanoscale HAR mesa of width $0.1\mu\text{m}$. Due to the generation of two closely superimposed edge waves the mesa appears as a straight line in the image, and the result agrees with 2D models (Figure 5.11).

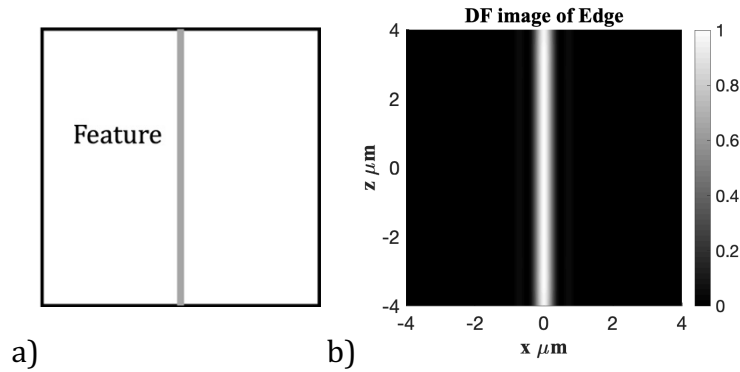


Figure 6.12: a) Layout and b) DF image of a nanoscale HAR mesa. TE polarization.

Similarly, Figure 6.13 shows TE images of a single mesa of different widths: a) $0.1\mu\text{m}$, b) $0.5\mu\text{m}$ and c) $1\mu\text{m}$. Figure 6.13a appears as a single bright line, Figure 6.13b shows a small valley between two bright lines and Figure 6.13c shows two distinct lines as the edge waves are separately resolved. Figure 6.13d quantifies these effects in terms of line scans along $z = 0$ for each image, as blue, red, and yellow lines respectively. Mesa edges can be identified as intensity peaks caused by the generation of edge waves. The peaks are larger when the edge waves are almost superimposed. This result again agrees with the investigation of 2D feature width in the previous chapter (Figure 5.19).

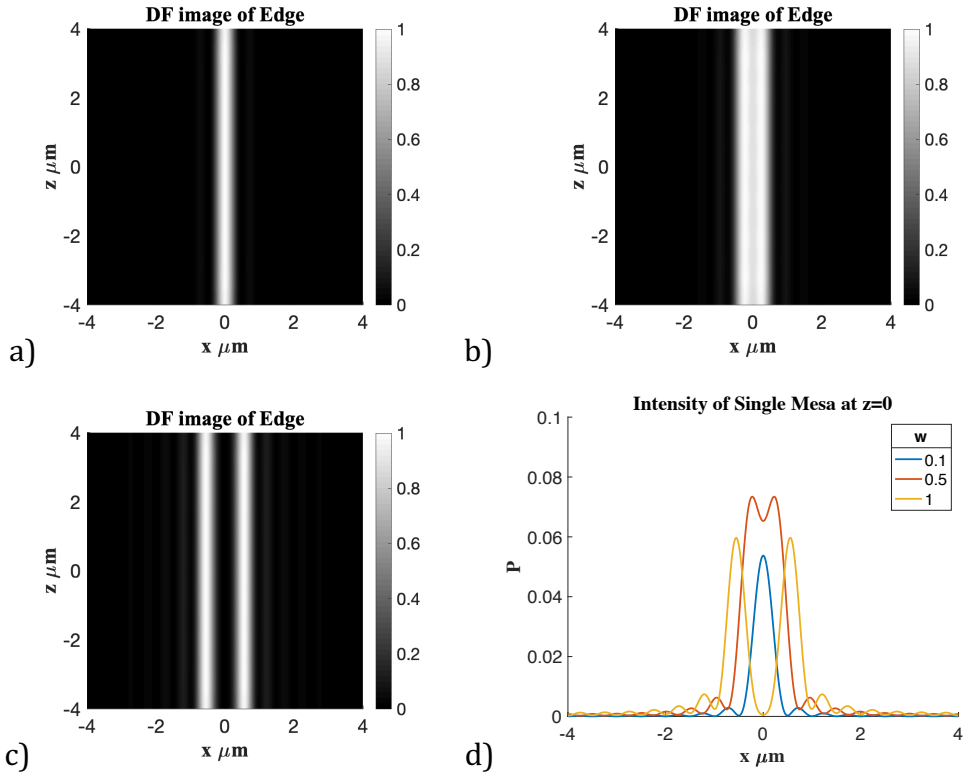


Figure 6.13: DF images of a mesa of width a) $0.1\mu\text{m}$, b) $0.5\mu\text{m}$ and c) $1\mu\text{m}$. d) Line scans along $z = 0$. TE polarization.

6.4.2 Double mesas

The second model is a pair of closely spaced, parallel mesas. Figure 6.14 shows TE images of a double mesa configuration with a separation of $1\mu\text{m}$ and widths of a) $0.01\mu\text{m}$ and b) $0.1\mu\text{m}$. The apparent widths and locations of the two bright lines are the same in both figures, and the images are analogous to that of a single mesa of $1\mu\text{m}$ width (Figure 6.13c).

This result suggests that highly similar images can be generated by structures with different dimensions, or by completely different structures. At first sight, this conclusion appears to violate a fundamental principle, namely the uniqueness of solutions to Maxwell's equations. However, it must be remembered that significant parts of the scattered field are discarded by the lens, and that only edge waves are contributing to the image; we shall present other examples of dissimilar object pairs that apparently look similar in a microscope later.

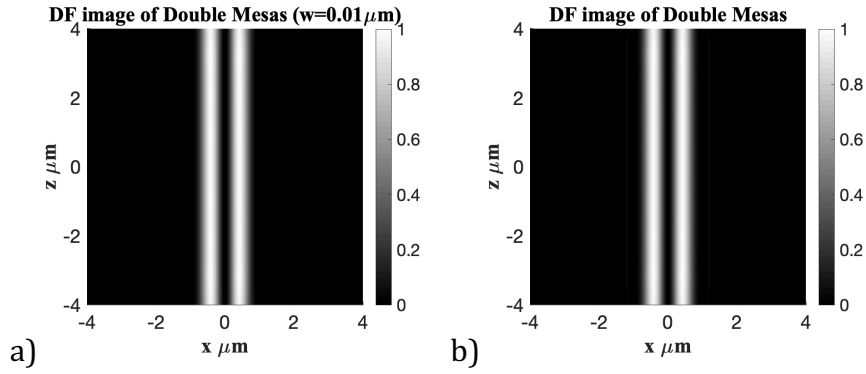


Figure 6.14: DF images of double mesas with $s = 1\mu\text{m}$ and widths of a) $0.01\mu\text{m}$ and b) $0.1\mu\text{m}$. TE polarization.

6.4.3 Crossed mesas

The third model consists of two mesas intersecting at a right angle, shown in Figure 6.15a. Figure 6.15b shows images for a mesa of width $0.1\mu\text{m}$. Here, the mesas appear as lines, while the intersection appears as a brighter, circular spot. The intersection area is $0.1\mu\text{m} \times 0.1\mu\text{m}$, small enough compared with the wavelength to act as a point source, so the intersection appears as an Airy disc.

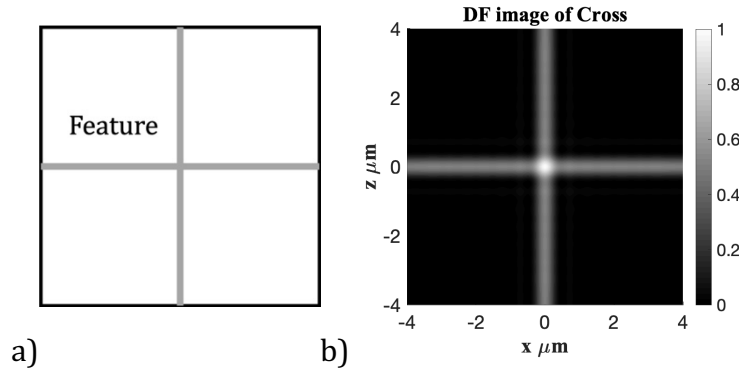


Figure 6.15: NEMS cross model: a) layout, and b) DF image. TE polarization.

Figure 6.16 shows a similar image for MEMS features of width $1\mu\text{m}$. Figure 6.16a shows the arrangement and 6.16b the DF image. In this image, the bright central feature has disappeared. Only the mesa edges can now be seen, and there is some edge rounding at the intersection point. This may be due to omission of corner detail caused by the limited number of illuminating waves.

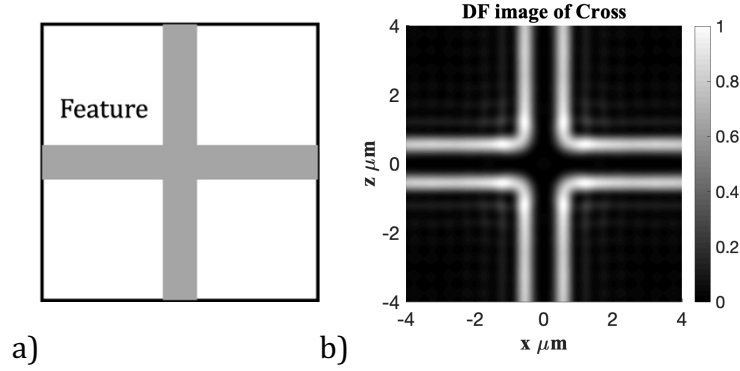


Figure 6.16: MEMS cross model: a) layout, and b) DF image. TE polarization.

6.4.4 Double crossed mesas

The fourth model consists of pairs of double mesas intersecting perpendicularly. Figure 6.17 shows a plan view of the configuration (6.17a) and compares TE DF images of a double mesa (6.17b) and a double cross configuration (6.17c), in each case for features of width $0.1\mu\text{m}$ and separation $1\mu\text{m}$. In both images, the bright lines have the same apparent widths and locations. However, the intersections are again brighter in the double cross image, as discussed above. In fact, the image of a double cross configuration is simply the incoherent sum of the images of a double mesa and its rotated version.

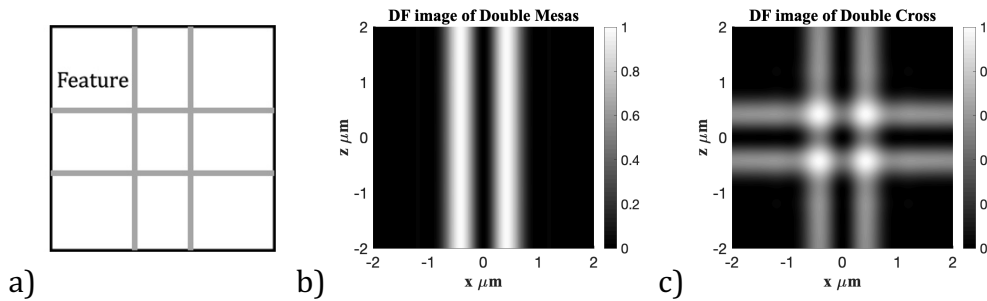


Figure 6.17: a) Double cross layout; DF images of b) double mesa and c) double cross. TE polarization.

6.4.5 Star patterns

The fifth model consists of the intersection of four mesas at 45° separation, forming a star pattern. Figure 6.18a shows a plan view of the structure. Figure 6.18b shows a TE image obtained using a mesa width $w = 0.1\mu\text{m}$. In contrast to crossed mesas, the brightness is now modulated in a radial direction; this may be due to multiple scattering effects.

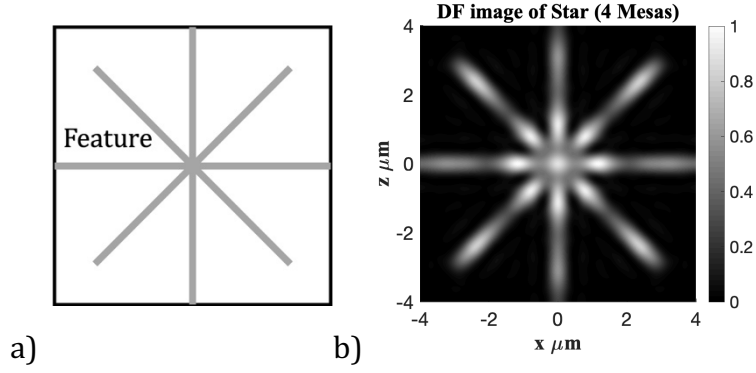


Figure 6.18: Star layout: a) layout, b) DF image. TE polarization.

6.4.6 Pillars and cylinders

Many other canonical structures have been simulated using FEM in the course of this work. We now consider other examples, starting with pillar structures to study the difference between solid and hollow sections. Figures 6.19a–d show the layouts of solid and hollow circular and square sections, and Figures 6.19e–h show the corresponding TE images. As before, only the feature outlines are visualised, as a circular ring or a square respectively. Here images are generated from eight incident waves to better describe the structure that is not in Manhattan arrangement. In general, the difference between the images of solid and hollow structures is negligible.

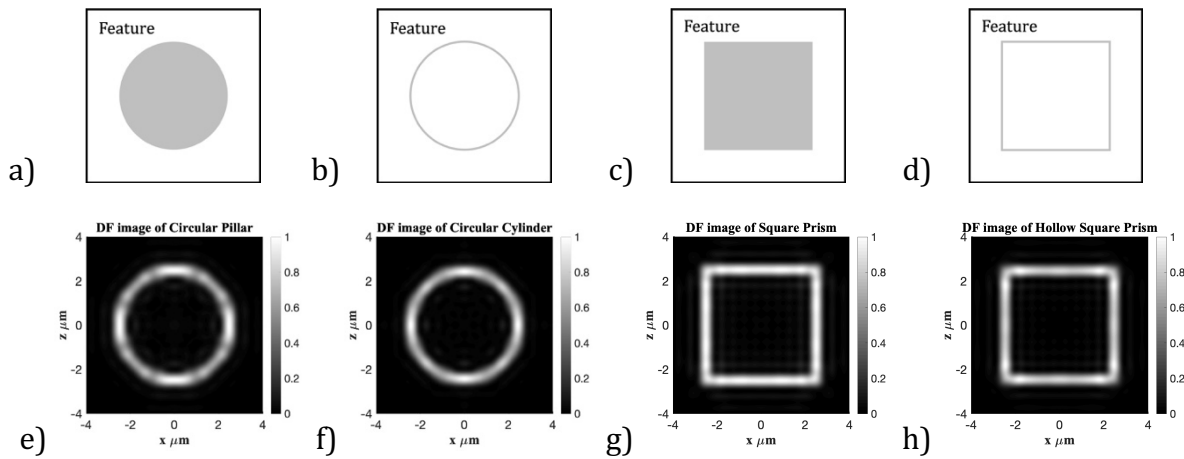


Figure 6.19: Layouts of a), b) solid and hollow cylinder, b), c) solid and hollow square prism. e)-h) corresponding images at $d = 1\mu m$. TE polarization.

Figures 6.20a-b show pillars with alternative rhomboid and elliptical cross sections, and Figures 6.20c–d show the corresponding TE DF images generated from eight illuminating

waves. These results again demonstrate the dependence of image brightness on outline curvature and highlight rapid changes at sharp corners or in regions of high curvature.

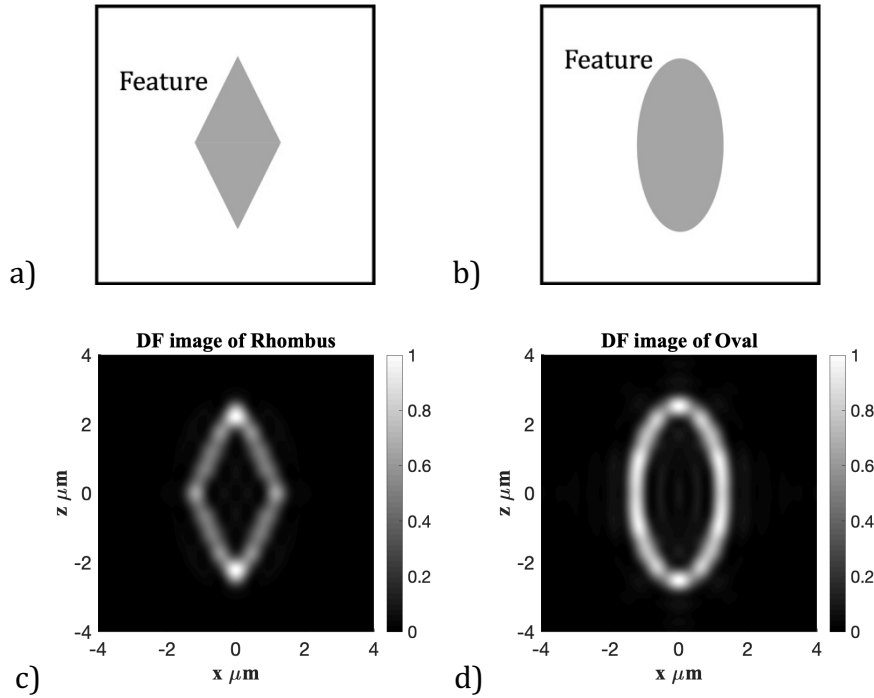


Figure 6.20: Layouts of a) a rhomboid and b) elliptical pillars; c)-d) corresponding DF images. TE polarization.

6.4.7 Finite structures

Now we consider other structures of finite length, such as incomplete mesas and crosses. Figure 6.21 shows the layout of a) a finite nanoscale mesa of $1\mu m$ length, b) a mesa with a $1\mu m$ gap and c) a cross with the intersection removed. In each case, the mesa width is $w = 0.1\mu m$. Figure 6.21d-f shows the corresponding TE DF images. Here, the main effect occurring is that the feature lengths appear shorter than the actual geometries. One possible explanation is that the terminating edge wave is spherical rather than cylindrical. In these regions, image brightness is therefore reduced, similar to the corners of a square prism. However, there is also some brightness modulating along the lengths of the mesas.

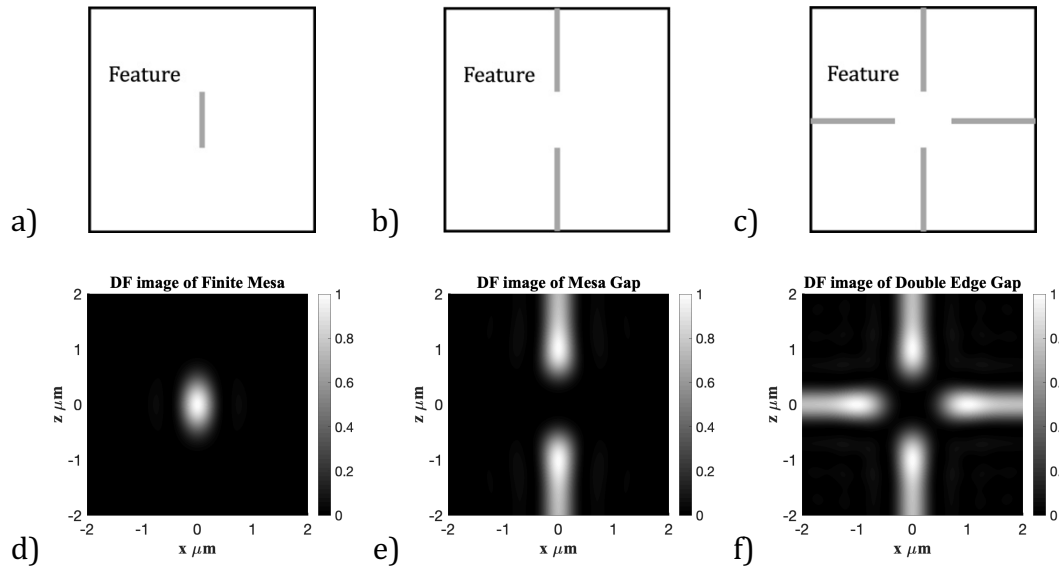


Figure 6.21: Layouts of a) mesa of finite length, b) mesa with a gap and c) incomplete mesa intersection. d)-f) corresponding DF images. TE polarization.

6.4.8 Indeterminate cases

We now consider additional several special cases involving different structures that yield apparently similar images. Previous examples have included single and double mesas (Figures 6.13 and 6.14). Here we consider two more. Figure 6.22 shows a nanoscale double cross (6.22a) and an incomplete MEMS cross, with the intersection removed (6.22b). Figures 6.22c and 6.22d show the corresponding TE DF images, which are almost identical. The main visible difference is the shape of the intersection points, due to the type of contact in each case. An additional periodic modulation in brightness can be seen in the central region of Figure 6.22d; its origin is unclear. This indeterminacy can be resolved numerically by plotting a linescan across the image and observe the difference in the intensity peak at the edge (referred to Figure 5.12b).

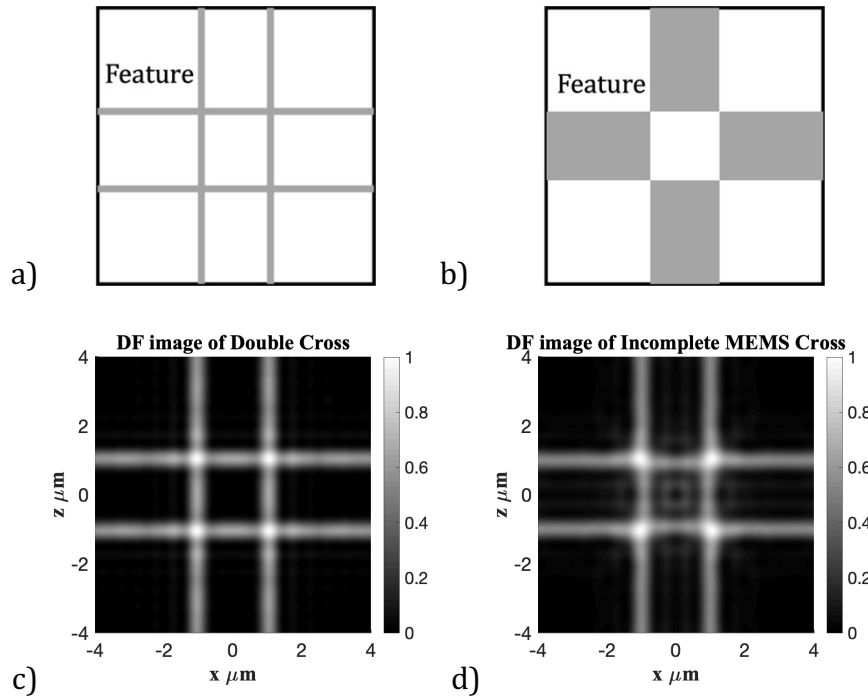


Figure 6.22: Layouts of a) double NEMS cross and b) incomplete MEMS cross. c) and d) corresponding images. TE polarization.

These examples confirm that DF image formation is dominated by edge waves, and that similar images will be obtained from objects that merely have the same edge layouts. It makes little difference whether these edges arise from nano- or micro-scale mesas, or whether the edges are internal or external. However, there are several methods to distinguish between the two structures in Figure 6.22a and 6.22b without processing the images.

Polarization

The first method to differentiate the two structures is to use other polarizations of illumination. Figure 6.23 shows the DF image of the double NEMS cross and the incomplete MEMS cross under TM polarization. Because of the significant secondary diffraction for small distance in TM polarization, the apparent width of the thin mesas is larger and the interception points become minor. In contrast, for the incomplete MEMS cross, the single edge waves of the microscale mesa appear identical to that in the TE image. Therefore, the difference between the two structures is easily observed by illuminating the structures in TM polarization or circular polarization.

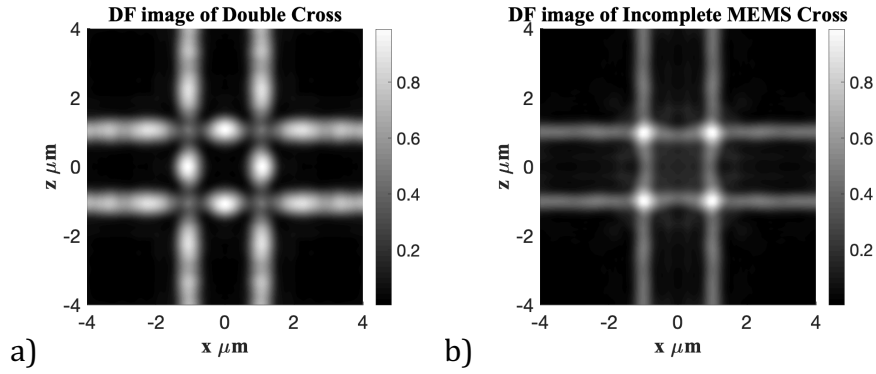


Figure 6.23: DF images of a) a double NEMS cross and b) an incomplete MEMS cross. TM polarization.

Bright Field

Another approach is to use the bright field illumination. Bright field microscope collects also the reflected field from the microscale mesa and produces a corresponding image. As seen in Figure 1.10a in Chapter 1, the difference between nano- and micro-scale is obvious in BF images. Figure 6.24 shows the BF image of the same set of structures at the lens angle $\theta_{max} = 60^\circ$ from monochromatic illumination at single incidence. The background is dark as no substrate is simulated as mentioned previously. The microscale mesa appears as bright rectangles in the image while the thin mesas in the double cross structure are seen as narrow bright lines. The structures in Figure 6.22a and 6.22b are pictured completely different in these BF images.

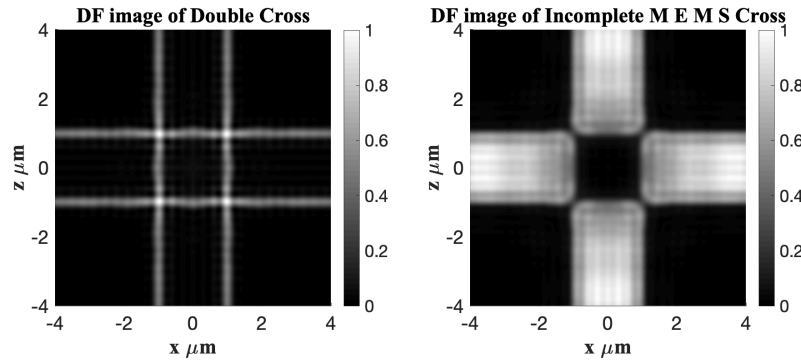


Figure 6.24: BF images of a) a double NEMS cross and b) an incomplete MEMS cross. TE polarization.

Confocal microscopy

Like the idea of the bright field illumination, confocal microscopy can be used to identify the structure between nano- and micro-scale. Using the confocal microscopy, the image

only captures the feature at the focal level. Due to the structural difference of the devices in Figure 6.22a and 6.22b, the images of each structure will appear differently.

Experimental flaw

Another possible method is to observe the dusts in the specimen in the microscope. For DF illumination, the dust will scatter off the incident light field and be captured by the objective lens. Since the dusts on the microscale mesa is at the imaging level, the images of the dusts (bright dots) will be clear. The dusts on the substrate (between the thin mesas) will however be out-of-focus and blurry.”

6.5 Additional out-of-plane illumination angles and wavelengths

We have so far only discussed images obtained using a single value of θ_{in} and monochromatic light. The primary reason is to reduce simulation time, and we now consider the effect of additional out-of-plane illumination angles and wavelengths.

6.5.1 Multiple out-of-plane illumination angles

For a single thin mesa, the only difference caused by a variation in out-of-plane incidence angle is a change in field strength. For multiple angles, the image must therefore simply be a sum of very similar intensity distributions. However, for layouts with multiple edges, the relative phases of waves scattered from different edges must be important. A double mesa can be used to illustrate this effect. Figure 6.23 shows TE DF images of a two mesas of width $0.1\mu m$ and separation $1\mu m$, obtained at $\lambda = 0.5\mu m$ using incident waves with angles θ_{in} of a) 40° , b) 50° , c) 60° , d) 70° and e) 80° , with azimuth angles ϕ_{in} spread over a range from 0° to 330° in steps 30° . Figure 6.23f shows the corresponding intensity linescans at $z = 0$. The images are qualitatively similar, but the apparent position of the intensity peaks clearly alters as θ_{in} varies. These differences are quantified in the linescans; those at 40° and 70° are approximately the same, but there are significant differences at other angles.

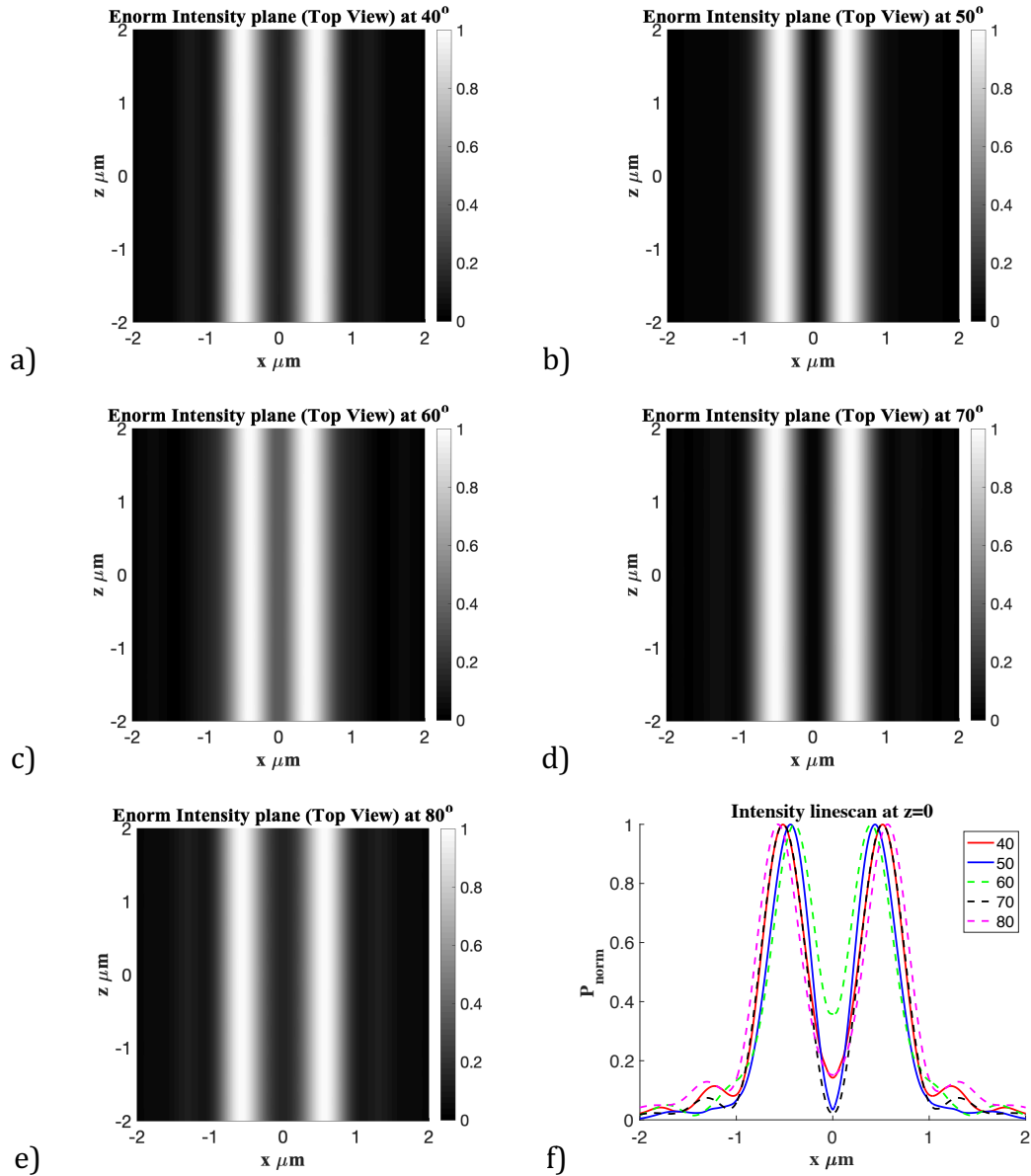


Figure 6.23: Simulated DF images of a double mesa using out-of-plane angles of a)-e) 40° to 80°; f) corresponding normalised intensity linescans along $z = 0$. TE polarization.

Despite the above, we would expect these details to blur out in an overall image obtained using all illumination waves together. Figure 6.24 shows the image obtained by summing contributions from all five incident angles θ_{in} in Figure 6.23. Note that the image here is almost exactly the same as the result obtained with a single wave incident at $\theta_{in} = 40^\circ$ (Figure 6.23a). This results suggest that the dominant contribution may come from illumination at angles lying closest to the acceptance cone of the lens.

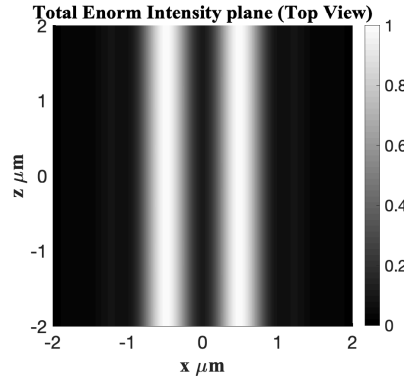


Figure 6.24: Full DF image of a double mesa. TE polarization.

6.5.2 Polychromatic illumination

In general, when the wavelength of the illumination increases or decreases, the apparent width of the intensity peak also increases or decreases, and the sidelobe pattern alters. To illustrate this, Figure 6.25 shows TE DF images of a single mesa of width $0.1\mu\text{m}$ at different wavelengths, namely a) $0.4\mu\text{m}$, b) $0.5\mu\text{m}$, c) $0.6\mu\text{m}$ and d) $0.7\mu\text{m}$. The bright line in the middle is clearly resolved less well as the wavelength increases.

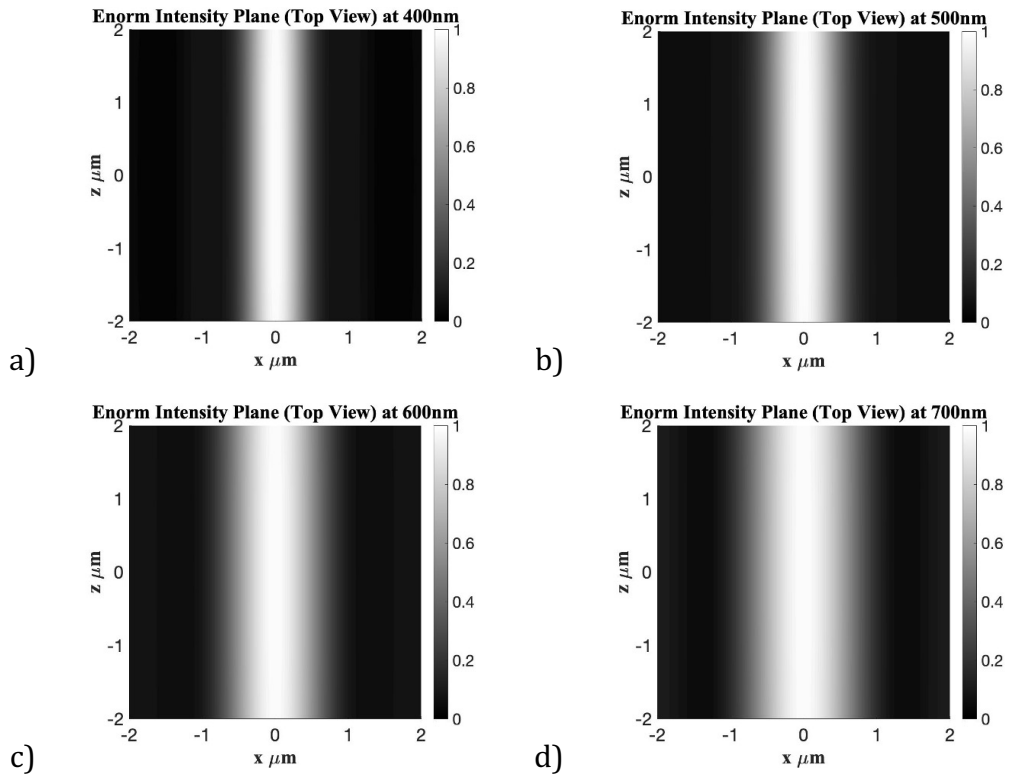


Figure 6.25: Simulated DF images of a single mesa for illumination at a wavelength of a) $0.4\mu\text{m}$, b) $0.5\mu\text{m}$, c) $0.6\mu\text{m}$ and d) $0.7\mu\text{m}$. TE polarization.

Again, we would expect these details to blur out under polychromatic illumination. Figure 6.26 shows the full DF image of a single mesa obtained by incoherently summing contributions at all four wavelengths shown in Figure 6.25, converting the result to black-and-white. Surprisingly, the apparent linewidth matches the linewidth obtained at $\lambda = 0.5\mu\text{m}$ alone, suggesting that this wavelength is dominant. However, the explanation is currently unclear, and extensive investigation of polychromatic illumination must therefore be considered to lie outside the scope of this thesis.

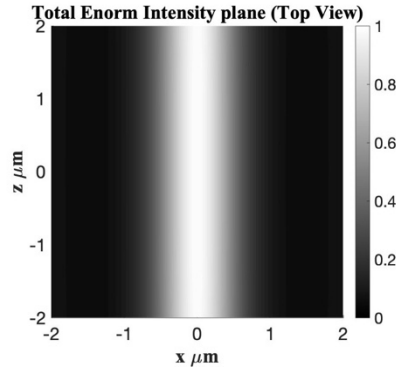


Figure 6.26: B/W equivalent polychromatic DF image of a single mesa. TE polarization.

6.6 TM polarization

In the previous sections, dark field monochromatic and polychromatic images of canonical structures were discussed in detail for TE polarization. We now present a smaller number of monochromatic examples for TM polarization for comparison.

Figure 6.27 shows normalised TM dark field images of single mesas of width a) $0.1\mu\text{m}$, b) $0.5\mu\text{m}$ and c) $1\mu\text{m}$ at $\theta = 50^\circ$. The images are very similar to the TE images in Figure 6.13. Figure 6.27d shows the line intensity variations in a) to c) along $z = 0$ as blue, red and yellow lines correspondingly. Unlike TE polarization, the image brightness obtained for the mesa of width of $w = 0.5\mu\text{m}$ is approximately twice that of the other two mesas, due to the secondary diffraction between the two edges. This result confirms not only that diffraction along the mesa plateau is important in TM polarization, but also that the secondary diffraction between two closely placed edges must also be significant.

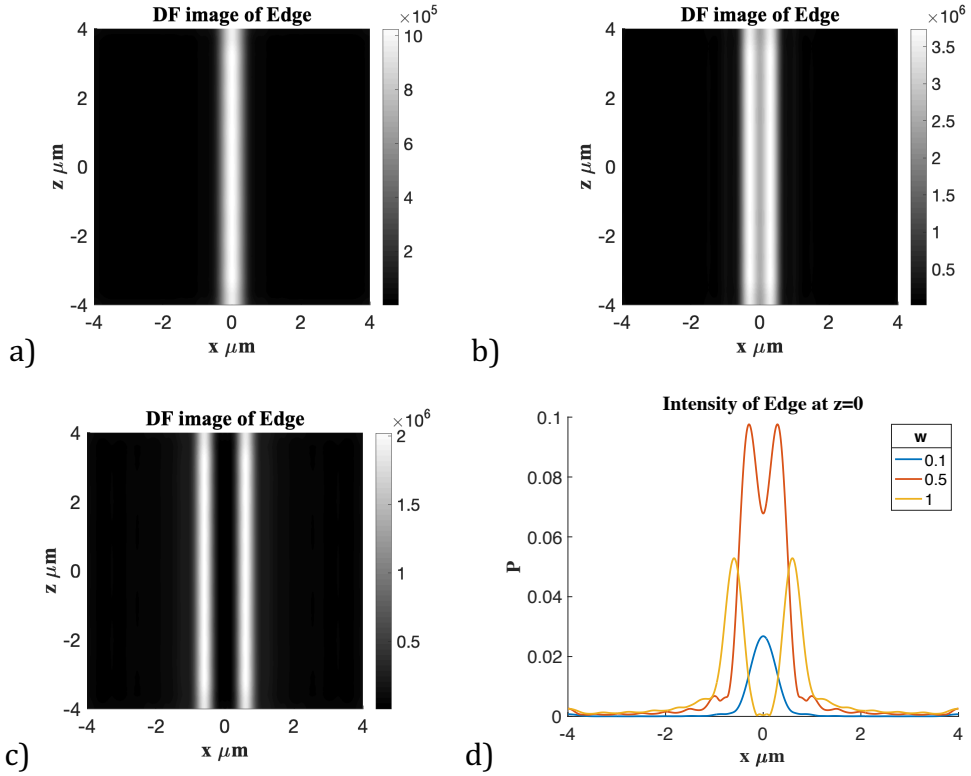


Figure 6.27: DF images of a mesa of width a) $0.1\mu\text{m}$, b) $0.5\mu\text{m}$ and c) $1\mu\text{m}$ in TM polarization. d) Line scans along $z = 0$. TM polarization.

Figure 6.28 shows the DF images of a) a square prism and b) a cylinder in TM polarization at $\theta = 50^\circ$. The images look highly similar to the TE images in Figure 6.7a and 6.11a. However, if we remove the normalisation of the DF images, the actual magnitude of the image fields in TM polarization is approximately 60% of the TE images, suggesting that TE polarization may be dominant in many cases.

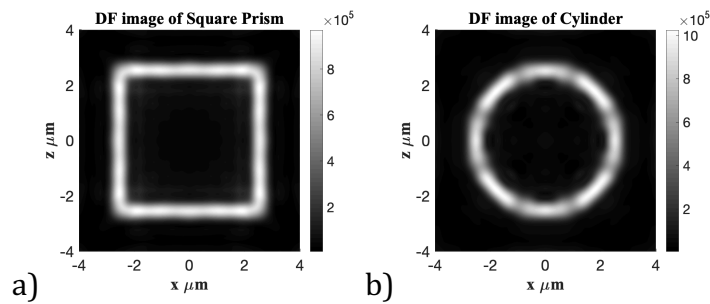


Figure 6.28: DF images of a) a square prism and b) a cylinder in TM polarization.

6.7 Summary

3D simulation of optical imaging has been validated by comparison with 2D models. The effect of illumination has been investigated, and its inherent selection of near perpendicular edges has been noted. Images of canonical structures have been simulated and found to represent a consistent but modified representation of the true geometry. The correlated appearance of solid and hollow features highlights the fact that images are formed by superposition of edge waves. For structures of finite length, there are discrepancies between DF images and the feature geometry, with the former showing dimensional reduction. Pairs of structures have been found to generate similar images, suggesting that imaging may not always provide unique results. The effect of additional out-of-plane illumination angles and wavelengths has been investigated and found to be small. Differences between images obtained for TE and TM polarization have been identified. In the next chapter, an analytical model of scattering based on diffraction coefficients will be developed from these observations to enable much more rapid simulation of dark field imaging.

References

3D Optical Microscopy

1. A. R. Stokes, "Three-dimensional diffraction theory of microscope image formation", *Proc. R. Soc. London, A*, vol. 212, pp. 64-274, 1952.
2. N. Streibl, "Three-dimensional imaging by a microscope", *JOSA A*, vol. 2, pp. 121-127, 1985.
3. C. J. R. Sheppard, X. Mao, "Three-dimensional imaging in a microscope", *JOSA*, vol. 6, pp. 1260-1269, 1989.
4. P. Török, P. R. T. Munro, Em. E. Kriezis, "High numerical aperture vectorial imaging in coherent optical microscopes", *Opt. Exp.*, vol. 16, pp. 507-523, 2008.
5. C. W. McCutchen, "Generalised aperture and the three-dimensional diffraction image", *JOSA*, vol. 54, pp. 240-244, 1964.
6. B. R. Frieden, "Optical transfer of a three-dimensional object", *JOSA*, vol. 57, pp. 56-66, 1967.
7. G. A. Laub, G. Lenz, E. R. Reinhardt, "Three-dimensional object representation in imaging systems", *Opt. Engng.*, vol. 24, pp. 901-905, 1985.
8. E. Evans, "Comparison of the diffraction theory of image formation with the three-dimensional, first Born scattering approximation in lens systems", *Opt. Comm.*, vol. 2, pp. 317-320, 1970.

3D Simulation of Scattering

9. G. L. Wojcik, D. K. Vaughan, L. K. Galbraith, "Calculation of light scatter from structures on silicon surfaces", *Proc. SPIE*, vol. 774, pp. 21-32, 1987.
10. G. Wojcik, J. Mould, R. Monteverde, J. Prochazka, J. Frank, "Numerical simulation of thick-linewidth measurements by reflected light", *Proc. SPIE*, vol. 1464, pp. 187-203, 1991.
11. A. K. Wong, A. R. Neureuther, "Rigorous three-dimensional time-domain finite-difference electromagnetic simulation for photolithographic applications", *IEEE Trans. Semicon. Manuf.*, vol. 8, pp. 419-431, 1995.
12. R. Guerrieri, T. K. Tadros, J. Gamelin, A. R. Neureuther, "Massively parallel algorithm for scattering in optical lithography", *IEEE Trans. Comp. Aid. Des. IC S.*, vol. 10, pp. 1091-1100, 1991.
13. I. R. Capoglu, C. A. White, J. D. Rogers, H. Subramanian, A. Taflove, V. Backman,

- “Numerical simulation of partially-coherent broadband optical imaging using the FDTD method”, *Opt. Lett.*, vol. 36, pp. 1596-1598, 2011.
14. M. Totzek, “Numerical simulation of high-NA quantitative polarization microscopy and corresponding near-fields”, *Optik*, vol. 112, pp. 399-406, 2001.
 15. J. Qin, R. M. Silver, B. M. Barnes, H. Zhou, R. G. Dixon, M.-A. Henn, “Deep subwavelength nanometric image reconstruction using Fourier domain optical normalisation” *Light: Sci. Apps.*, vol. 5, pp. 1-9, 2016.
 16. H. Tanabe, “Modeling of optical images in resists using vector potentials”, *Proc. SPIE*, vol. 1674, pp. 637-649, 1992.
 17. K. D. Lucas, H. Tanabe, A. J. Strojwas, “Efficient and rigorous three-dimensional model for optical lithography simulation”, *JOSA A*, vol. 13, pp. 2187-2199, 1996.
 18. J. Kim, Y. Wang, X. Zhang, “Calculation of vectorial diffraction in optical systems”, *JOSA A*, vol. 35, pp. 526-535, 2018.
 19. L. E. R. Petersson, G. S. Smith, “Three-dimensional electromagnetic diffraction of a Gaussian beam by a perfectly conducting half-plane”, *J. Opt. Soc. Am.*, vol. 19, pp. 2265-2280, 2002.
 20. E. Marx, “Integral equations for 3D scattering: finite strip on a substrate”, *PIERS Online*, vol. 5, pp. 133-136, 2009.
- 3D Scattering Theory
21. B. Richards, E. Wolf, “Electromagnetic diffraction in optical systems II. Structure of the image field in an aplanatic system”, *Proc. R. Soc. London., A*, vol. 253, pp. 358-379, 1959.
 22. J. B. Keller, “Geometrical theory of diffraction”, *J. Opt. Soc. Am.*, vol. 52, pp. 116-130, 1960.
 23. R. R. A. Syms, A. Bouchaala, O. Sydoruk, D. Liu, “Optical imaging and image analysis for high aspect ratio NEMS,” *J. Micromech. Microeng.*, vol. 29, pp. 1-11, 2018.
 24. C. A. Balanis, “Radiation from slots on cylindrical bodies using geometrical theory of diffraction and creeping wave theory”, *NASA TR R-331*, Washington D. C., 1970.
 25. H. Ikuno, M. Nishimoto, “Calculation of transfer functions of three-dimensional indented objects using the physical optics approximation combined with the method of stationary space”, *IEEE Trans. Ant. Propag.*, vol. 39, pp. 585-590, 1991.
 26. K. Belkebir, P. C. Chaumet, A. Sentenac, “Influence of multiple scattering on three-dimensional imaging with optical diffraction tomography”, *JOSA A*, vol. 23, pp. 586-

595, 2006.

27. G. Carluccio, M. Albani, "An efficient ray tracing algorithm for multiple straight wedge diffraction", IEEE Trans. Ant. Propag., vol. 56, pp. 3534-3542, 2008.
28. G. Carluccio, F. Puggelli, M. Albani, "Generalization of UTD double diffraction to the case of impenetrable wedges with relatively general boundary conditions", IEEE Trans. Ant. Propag., vol. 62, pp. 3829-3834, 2014.

3D Imaging

29. I. R. Capoglu, J. D. Rogers, A. Taflove, V. Backman, "The microscope in a computer: image synthesis from three-dimensional full-vector solutions of Maxwell's equations at the nanometer scale", Prog. Opt., vol. 57, pp. 1-91, 2012.
30. B. M. Barnes, M. Y. Sohn, F. Goasmat, H. Zhou, A. E. Vladár *et. al.*, "Three-dimensional deep sub-wavelength defect detection using $\lambda = 190nm$ optical microscopy", Opt. Exp., vol. 21, pp. 26219-26226, 2013.
31. R. Attota, R. G. Dixon, "Resolving three-dimensional shape of sub-50nm wide lines with nanometer-scale sensitivity using conventional optical microscopes", Appl. Phys. Lett., vol. 105, pp. 043101, 2014.
32. Z. Liu, L. Tian, S. Liu, L. Waller, "Real-time brightfield, darkfield and phase contrast imaging in a light-emitting diode array microscope", J. Biomed. Optics, vol. 19, pp. 106002, 2014.

7. Novel Analytical Simulation Tool

This chapter will attempt to develop simple scalar models of dark field microscopy for 2D and 3D HAR object geometries, using observations from the previous chapter to simplify modelling of illumination, the geometric theory of diffraction to provide diffraction coefficients, and Fourier transformation, spatial filtering, and inverse transformation to describe image formation. The structures modelled in 2D will be restricted for both TE and TM incidence. However, the results will be used to develop a model for arbitrary structures in 3D. The main differences between the models are the additional illuminating waves and the use of ITD and Huygens' principle to model in-plane scattering. In all cases, model predictions will be bench-marked against 2D and 3D FEM.

7.1 Introduction

Rigorous scattering from isolated and periodic 2D sub-wavelength structures has already been reviewed, but the discussion focused on primary diffraction. For closely placed features, re-scattering or secondary diffraction is significant and is reviewed here in more detail. Diffraction from multiple knife-edges was studied in the early 1960s, focusing on transmission loss over knife-edge terrain obstacles [1-3]. For example, Figure 7.1 shows a ray tracing model of signal propagation from the transmitter T to the receiver R over hills H1 and H2 [4, 5]. Vogler's solution [6], based on the work of Furutsu [7], starts with a generalised residue series for propagation over obstacles, obtained by a Green's function formulation. Other approaches were developed using the Fresnel-Kirchhoff theory of scalar waves [8, 9], Wiener-Hopf techniques [10] and the UTD [11, 12].

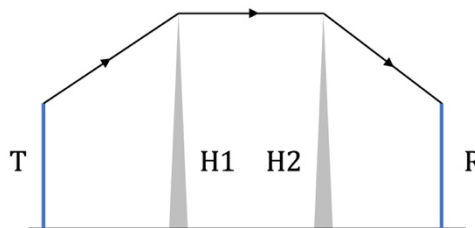


Figure 7.1: Double knife edge diffraction of terrestrial propagation.

Since the multiple half-plane model is inappropriate for our own diffraction scenarios, we now review diffraction from multiple wedges. One of the earliest studies formulates the scattering from perfectly conducting body with an iterative surface current density replacement technique [13]. It calculates the diffraction from closely spaced edges

accurately and has been verified by GTD. However, the difficulty is the integration. Other studies of multiple wedge diffraction by Tiberio [14-16], Elsherbini [17] and Michaeli [18, 19] are restricted by their own limitations. A simple and exact expression of double diffraction can be formulated in a similar fashion to UTD [20, 21], as:

$$E_{dd} = E_0 D_1 D_2 \frac{e^{-jkr}}{\sqrt{rr'}} \quad (7.1)$$

Here D_1 and D_2 are the diffraction coefficients of the wedges Q_1 and Q_2 . Figure 7.2 shows the geometry. ϕ and ϕ' are the incident angles at each wedge and ϕ_1 and ϕ_2 are the diffracted angles. Primary diffraction from each wedge is calculated using the GTD expression (Equation 3.31), and the total field is found as the sum of the incident field, primary diffraction, and re-scattering between wedges.

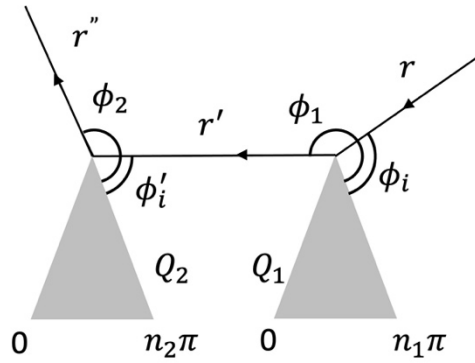


Figure 7.2: Geometry of double wedge diffraction.

Extension to 3D can be achieved by considering the azimuth angles in scattering between the two wedges [22-25]. For multiple wedges, the formulation is often based on matrices [26, 27]. However, the significance of higher order scattering falls quickly due to the spreading factor. ITD is often used for calculation [28-30], and allows the diffracted field to be computed from the sum of incremental field contributions, obtained using their respective GTD diffraction coefficients [31, 32]. Recall the previous ITD equations:

$$E_d(P) = \int_L dE_d(P) dl = \int_L D(l) E_0(l) \frac{e^{-jkr(l)}}{r(l)} dl \quad (7.2)$$

Here P is the observation point and L is the edge under consideration. $dE_d(P)$ is the incremental field contribution along L (Equation 3.58). ITD is useful for 3D scattering problems as it simplifies scattering from complicated structures such as mesa ends.

Huygens' principle can also be invoked to find the elementary contribution of each point on a wavefront. This states that every point acts as a source of spherical wavelets and the

sum of the spherical wavelets gives the total wavefront [33]. A rigorous scalar wave theory of diffraction based on Huygens' principle has been formulated [34] and widely applied in wave propagation [35, 36] and modelling [37, 38].

Finally, we consider an image formation technique based on GTD for 2D structures, known as the diffraction edge response (DER) [39-41], which uses a heuristic relation between the edge of the structure and the disturbance in image space. For example, as we showed the previous chapter, scattering falls near feature ends. A modulated function of the square root of the intensity of the edge segment can then be used to form an image. The result is only approximate but can be computed rapidly. The motivation was optical proximity correction (OPC) and the study of mask alignment [42, 43]. A PTD based technique for OPC using edge diffraction was also developed but involves complex surface current integration and will not be considered here.

7.2 2D model – TE polarization

We now construct a scattering model for DF imaging of 2D objects. The aim will be to determine the minimum number of edge waves capable of modelling scattering from HAR objects, and hence to guide development of a 3D model. As before, we assume PEC surfaces, ignore the substrate, and omit image inversion. For a single edge, we proceed as follows. We assume that the only waves entering the lens are cylindrical edge waves:

$$E(r) = E_0 D \frac{\exp(-jk_0 r)}{\sqrt{r}} \quad (7.3)$$

Here r is the radial distance from the edge, $k_0 = 2\pi/\lambda$, and D is the diffraction coefficient. The diffraction coefficient will of course vary with the illumination and observation angles ϕ_0 and ϕ and the polarisation, and these variations are different in GTD, PTD and UTD. However, as shown before, these differences are unimportant away from the boundaries between regions in the space around the edge. In GTD, which we have previously shown to be a reasonable approximation, the coefficient for a wedge is:

$$D(\phi, \phi_0) = \frac{\exp\left(-j\frac{\pi}{4}\right)}{\sqrt{2\pi k_0}} d(\phi, \phi_0) \quad (7.4)$$

Here the angular dependency is specified by the reduced coefficient d . For *TE* incidence:

$$d = \frac{\sin\left(\frac{\pi}{n}\right)}{n} \left\{ \frac{1}{\cos\left(\frac{\pi}{n}\right) - \cos\left(\frac{\phi - \phi_0}{n}\right)} - \frac{1}{\cos\left(\frac{\pi}{n}\right) - \cos\left(\frac{\phi + \phi_0}{n}\right)} \right\} \quad (7.5)$$

Coefficient for *TM* incidence is calculated by replacing the minus sign with plus. Previous results have also shown that the variations with observation angle are slow and are approximately constant in the direction of the imaging lens. Assuming that expression in Equation 7.5 is D_1 for illumination of a leading edge, we may therefore write:

$$E \approx E_0 D_1 \frac{\exp(-jk_0 r)}{\sqrt{r}} \quad (7.6)$$

Here E_0 is the illuminating field. This pattern has a uniform distribution in k -space. However, the effect of the lens will be to limit the range of spatial frequencies to $\pm k_{max} = k_0 \sin(\theta_{max})$, where θ_{max} is the half angle of the lens. The effect can be found by the standard route of Fourier transformation, filtering by the pupil function, and inverse transformation. In this case, the object is so simple that the process may be carried out analytically. Ignoring the effect of the obliquity factor (which varies slowly), the electric field describing the image may then be found by inverse Fourier transformation, as:

$$E_{im}(x) = E_0' d_1 \int_{-k_{max}}^{k_{max}} \exp(jkx) dk = E_0'' d_1 \text{sinc}(k_{max}x) \quad (7.7)$$

Here E_0' and E_0'' are constants. The corresponding intensity image $I_{im} = |E_{im}|^2$ is:

$$I_{im} = |E_0'' d_1|^2 \text{sinc}^2(k_{max}x) \quad (7.8)$$

The phase of the illuminating wave can be included by making E_0'' complex. However, it should be noted that Equation 7.8 is unchanged in the process. Thus, the image of a single edge will always approximate a sinc^2 function, whose width is determined by k_{max} , so the resolution will (as expected) be controlled entirely by λ and θ_{max} .

Displaced edges will similarly generate displaced images. For example, the amplitude image of an edge located at $x = \delta$ is a shifted sinc function:

$$E_{im} = E_0'' d_1 \text{sinc}\{k_{max}(x - \delta)\} \quad (7.9)$$

Images of multiple edges may be computed as coherent sums. For example, the amplitude image of two edges at δ_1 and δ_2 with illuminations E_{01}'' and E_{02}'' and diffraction coefficients d_1 and d_2 is:

$$E_{im} = E_{01}'' d_1 \text{sinc}\{k_{max}(x - \delta_1)\} + E_{02}'' d_2 \text{sinc}\{k_{max}(x - \delta_2)\} \quad (7.10)$$

In this case, the effect of illumination phase will be significant, because cross terms will be generated in calculating the intensity. Consequently, both the direction and the

wavelength of the illuminating beam will affect the image. However, if $|\delta_1 - \delta_2|$ is large enough, differences will appear mainly in the sidelobes between the two edge images.

These observations may be used to develop scattering models for sets of edge objects derived from mesas, trenches, and closely spaced mesas as shown in Figure 7.3. Here reduced diffraction coefficients $d_1, d_2 \dots d_5$ corresponding to the coefficients $D_1, D_2 \dots D_5$ may be obtained from GTD (and have been computed earlier).

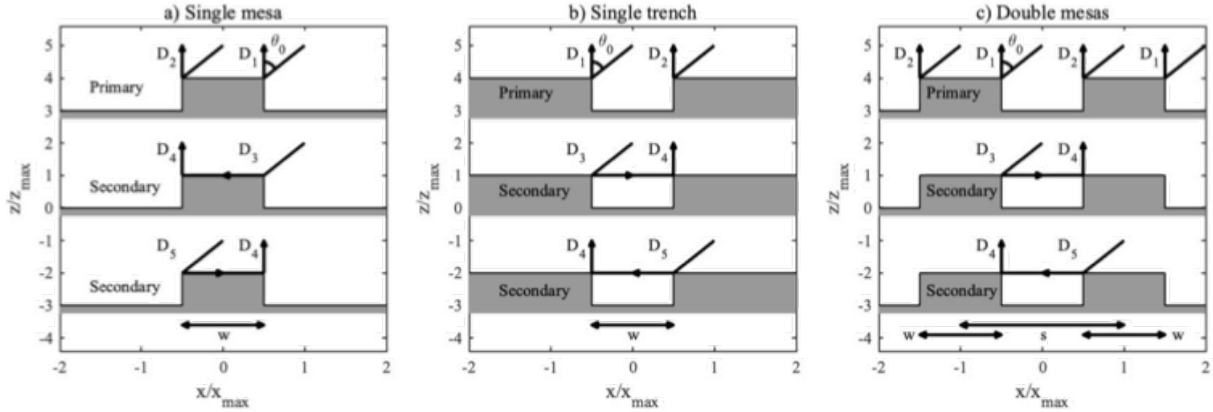


Figure 7.3: Scattering from a) single mesas, b) single trenches and c) double mesas.

We begin by considering scattering from a single mesa of width w , shown in Figure 7.3a, at TE incidence. We assume that the illuminating wave is a single wave travelling obliquely at an angle θ_0 , and hence described on the object plane by the field:

$$E_i = E_0 \exp\{+jk_0 x \sin(\theta_0)\} \quad (7.11)$$

In this case, we might expect primary scattering from the leading and trailing edges, described by diffraction coefficients D_1 and D_2 . We might also expect secondary scattering described by paths involving coefficients $D_3 \dots D_5$. However, their effect appears to be negligible (as does that of higher order scatterings) as we show later.

We therefore start by assuming that the overall electric field describing the image can be written as a coherent sum of images from the two edges:

$$E_{im}(x) = E_0'' \left\{ d_1 \exp\left\{+jk_0 \frac{w}{2} \sin(\theta_0)\right\} \text{sinc}\left[k_{max}\left(x - \frac{w}{2}\right)\right] + d_2 \exp\left\{-jk_0 \frac{w}{2} \sin(\theta_0)\right\} \text{sinc}\left[k_{max}\left(x + \frac{w}{2}\right)\right] \right\} \quad (7.12)$$

Here E_0'' is again an unimportant constant, and the first and second terms in the brackets are the images from the leading and trailing edges, respectively.

Figure 7.4 compares the prediction of this model (red) for the line variation of electric field modulus with corresponding results from 2D FEM (blue). Results are shown for mesa widths of $w = 0.01, 0.1, 0.2, 0.5, 1.0, 1.5, 2, 2.5$ and $3 \mu m$. In each case, the following parameters are assumed: $\lambda = 0.5 \mu m$, $\theta_{max} = 30^\circ$ and $\theta_0 = 50^\circ$. For FEM, the model space is $\Lambda = 8 \mu m$. For the scattering model, the diffraction coefficients were obtained from GTD; their values (and others needed later) are shown in Table 7.1. To reduce the number of coefficients, it is useful to work with a coefficient ratio $r_1 = d_2/d_1$, whose value may be found from Equation 7.5 as $r_1 = -0.3171$. The constant E_0'' is a fitting coefficient.

d_1	d_2	d_3	d_4	d_5
-1.8990	+0.6021	-2.1352	-1.7321	-3.0276
$R_1 = d_2/d_1$		$R_1 = +0.6021/-1.8990 = -0.3171$		
$R_2 = d_3 d_4/d_1$		$R_2 = -2.1352 \times -1.7321/-1.8990 = -1.9475$		
$R_3 = d_5 d_4/d_1$		$R_3 = -3.0276 \times -1.7321/-1.8990 = -2.7614$		

Table 7.1. TE GTD diffraction coefficients and ratios, for incidence at 50° to vertical.

Due to the single-sided illumination, the plots are asymmetric. For small w , mesa edges are unresolved and only a single peak can be seen. As w increases, two peaks appear. Each generates a sinc-function image, but scattering from the leading edge is clearly larger. Agreement between the models is excellent for all mesa widths. Additional numerical experiments were performed to investigate alternative scattering models. However, it was found that good agreement was obtained with the coefficient values predicted by GTD (and only those), and that the addition of the higher-order scattering terms shown in Figure 7.3a did not significantly improve agreement.

Figure 7.5 shows the corresponding normalised intensity variations, now assuming illumination by two waves at incidence angles $\pm\theta_0$ and computing the result as an incoherent sum. Because of the double-sided illumination, the plots are now symmetric; each edge that is leading for one illuminating wave becomes trailing for the other. For small w , the mesa is visualised as a single intensity peak. As w increases, two peaks appear, one at each edge. Agreement between the scattering model and FEM is again excellent, even at the sidelobe level. A small discrepancy occurs at $w = 0.5 \mu m$; since this dimension matches the wavelength, the effect is almost certainly due to resonant in-plane

scattering. However, the effect is relatively small, implying that the amplitude of any secondary scattering over the mesa plateau itself must also be small.

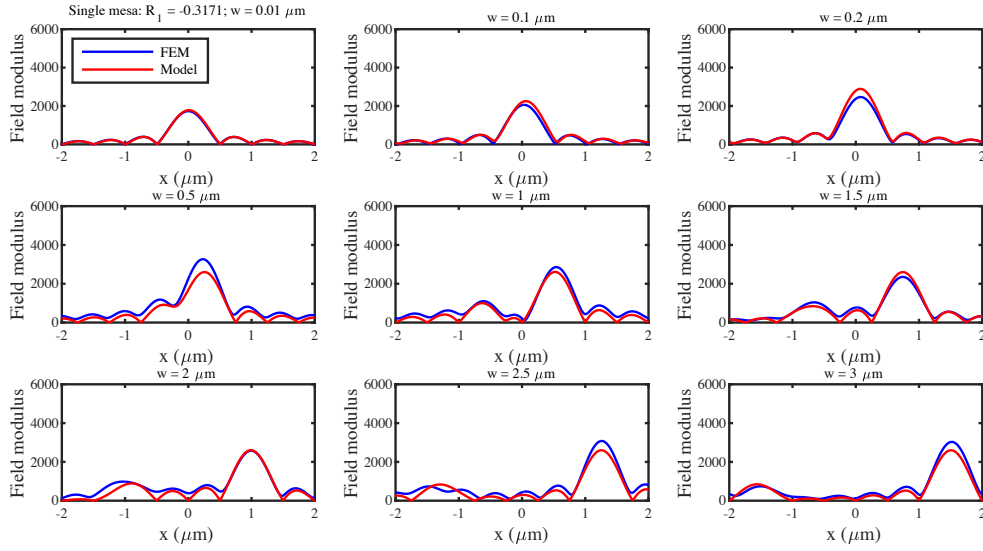


Figure 7.4: Single mesas - comparison of absolute line amplitude variations predicted by FEM and a model containing primary scattering only. TE polarization.

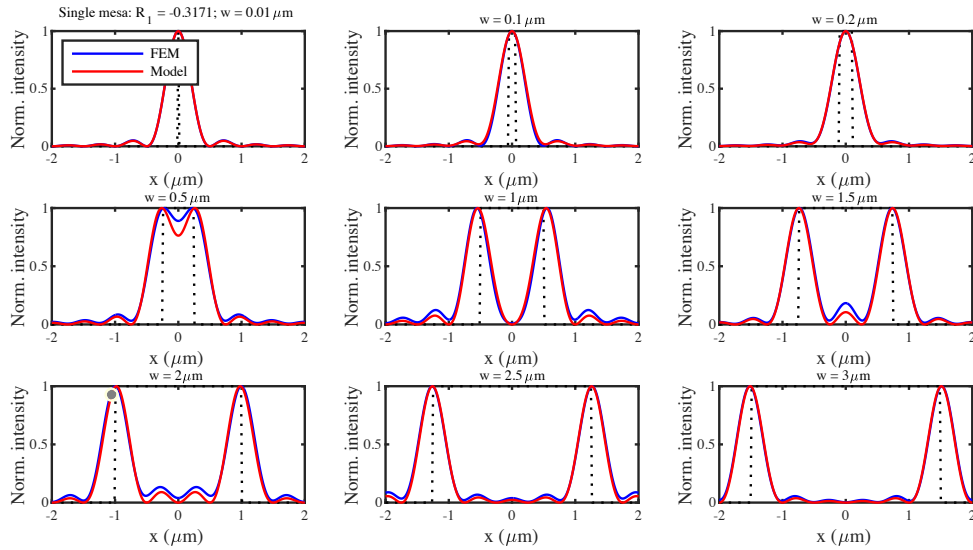


Figure 7.5: Single mesas - comparison of normalised line intensity variations predicted by FEM and a model containing primary scattering only. TE polarization.

We now consider scattering from a single trench, shown in Figure 7.3b. Again, we might expect primary scattering from the leading and trailing edges, described by diffraction coefficients D_1 and D_2 , and secondary scattering described by ray paths involving coefficients $D_3 \dots D_5$. Note that the directions of the re-scattered waves are different to

those in Figure 7.3a since the ray paths are now in air. In contrast to the mesa, numerical experiments have shown that these terms cannot be ignored. Higher order terms may also be important; however, to retain sufficient simplicity for a 3D model, we ignore these for now. We therefore assume that the electric field describing the line image of a single trench is:

$$\begin{aligned}
 E_{im}(x) = E_0'' \bigg\{ & d_2 \exp \left\{ +jk_0 \frac{w}{2} \sin(\theta_0) \right\} \text{sinc} \left[k_{max} \left(x - \frac{w}{2} \right) \right] \\
 & + d_1 \exp \left\{ -jk_0 \frac{w}{2} \sin(\theta_0) \right\} \text{sinc} \left[k_{max} \left(x + \frac{w}{2} \right) \right] \\
 & + d_3 f d_4 \exp \left\{ -jk_0 \frac{w}{2} \sin(\theta_0) \right\} \text{sinc} \left[k_{max} \left(x - \frac{w}{2} \right) \right] \\
 & + d_5 f d_4 \exp \left\{ +jk_0 \frac{w}{2} \sin(\theta_0) \right\} \text{sinc} \left[k_{max} \left(x + \frac{w}{2} \right) \right] \bigg\}
 \end{aligned} \quad (7.13)$$

Here the term f , which describes the in-plane radiation of a cylindrical wave, is:

$$f = \exp \left(-j \frac{\pi}{4} \right) \frac{\exp(-jk_0 w)}{\sqrt{2\pi k_0 w}} \quad (7.14)$$

Figure 7.6 compares the prediction of this model for the line variation of electric field modulus with corresponding results from 2D FEM. Results are shown for trench widths $w = 0.01, 0.1, 0.2, 0.5, 1.0, 1.5, 2, 2.5$ and $3 \mu\text{m}$. The main optical parameters are again as given before, and the diffraction coefficients are now defined by ratios $r_1 = d_2/d_1$, $r_2 = d_3 d_4/d_1$ and $r_3 = d_5 d_4/d_1$. Their values were obtained from GTD as $r_1 = -0.3171$, $r_2 = -1.9475$, and $r_3 = -2.764$ as shown in Table 7.1. E_0'' is the same fitting coefficient.

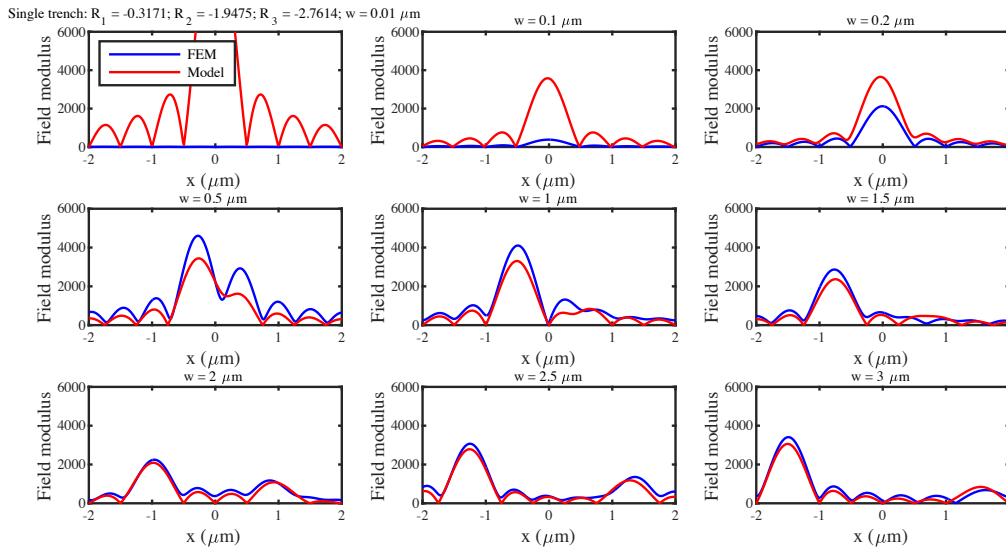


Figure 7.6: Single trenches - comparison of absolute line amplitude variations predicted by FEM and a model containing both primary and secondary scattering. TE polarization.

Once again, the variations are asymmetric, and the leading-edge image (which has now changed sides) is brighter. Good agreement with FEM is obtained for trench widths $w > 1 \mu\text{m}$. However, for narrow trenches, agreement is lost. Due to the $1/\sqrt{w}$ term in the expression for f , the simple scattering model predicts infinite fields as $w \rightarrow 0$. In contrast, FEM predicts that the scattering eventually disappears as the trench vanishes, consistent with physical expectations. This discrepancy implies that the expressions above cannot accurately model very small feature separations.

To demonstrate the need for higher terms, Figure 7.7 compares the prediction of the scattering model for the line variation of electric field modulus with corresponding results from 2D FEM, now omitting secondary scattering by setting $R_2 = R_3 = 0$. Agreement with FEM is clearly worse for trench widths in the range $0.5 \mu\text{m} \leq w \leq 1.5 \mu\text{m}$, when secondary scattering might be expected to be important. Further investigations have shown that inclusion of higher-order scattering terms can improve matters, but (since the additional terms also contain inverse powers of w) the results again diverge for narrow trenches. Given the additional complexity these terms will introduce to a 3D model, their inclusion is of marginal benefit.

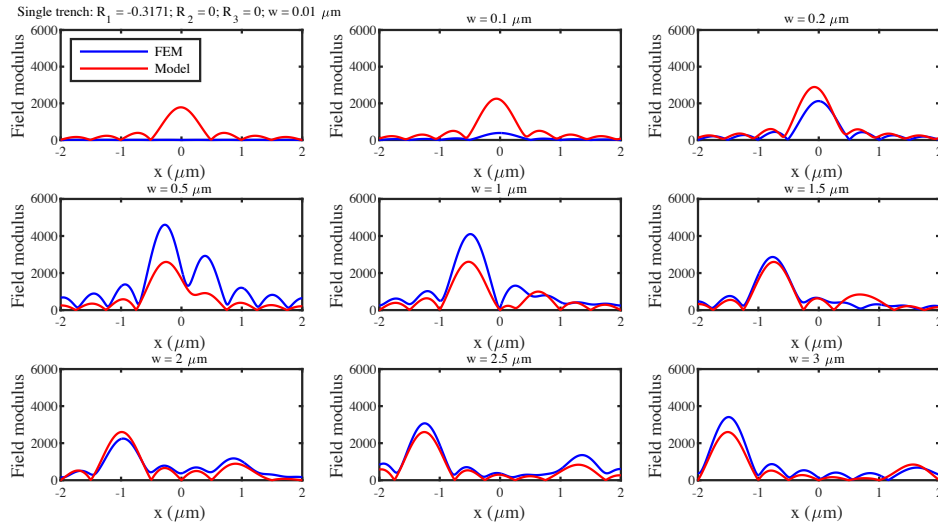


Figure 7.7: Single trenches - comparison of absolute line amplitude variations predicted by FEM and containing primary scattering only. TE polarization.

Figure 7.8 shows the normalised intensity variations corresponding to Figure 7.6, now assuming illumination by two waves at incidence angles $\pm\theta_0$ and computing the result as

an incoherent sum. Once again, the use of double-sided illumination has made the image symmetric. A resonance effect can again be noted for $w = 0.5\mu\text{m}$. Despite apparent agreement between the two models for $w = 0.2\mu\text{m}$, predictions of the scattering model for trench widths significantly smaller than $1.0\mu\text{m}$ may be regarded as inaccurate. For such tiny gap trench structures, the secondary diffraction falls into the near field range. Based on the “far field” assumption of GTD (mentioned in Chapter 3.2), the diffraction coefficients calculated in Table 7.1 do not accurately describe the secondary diffraction of small trench width structure.

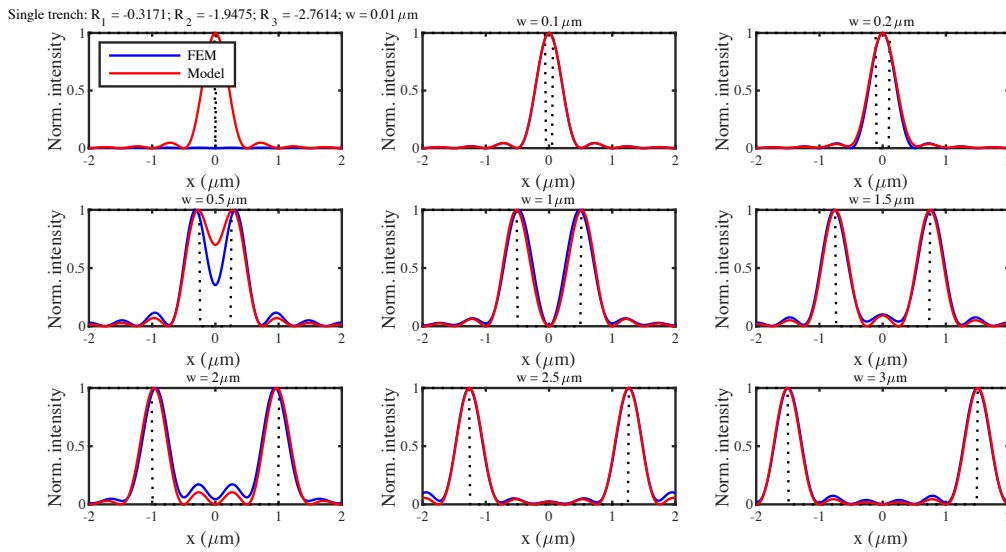


Figure 7.8: Single trenches - comparison of normalised line intensity variations predicted by the scattering model and FEM. TE polarization.

Finally, we consider scattering from a double mesa, shown in Figure 7.3c. Clearly, the scattering terms will reduce to those of isolated mesas when the trench width is large, and those of isolated trenches when the mesa width is large, and it trivial to construct the image field by combining Equations 7.12 and 7.13. Figure 7.9 compares the resulting single-sided amplitude and double-sided intensity images with FEM for mesa widths $w = 0.1\mu\text{m}$ and mesa separations of $s = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.5, 2$ and $2.5\mu\text{m}$, with all other parameters being as before. Except for small separations, agreement between the scattering model and FEM is excellent. There is significant disagreement at $s = 0.6\mu\text{m}$; this corresponds to an air gap of $0.5\mu\text{m}$ and in-plane resonance.

It should be emphasised that all the results in Figures 7.4, 7.5, 7.6, 7.7 and 7.8 were obtained using just three coefficient ratios, whose values were derived from fundamental theoretical considerations rather than a numerical matching process. Furthermore, good agreement with FEM was obtained over a range of structures and dimensions, merely provided features were not too closely spaced. Although approximate, the scattering model therefore appears both justifiable and sufficiently accurate for many purposes.

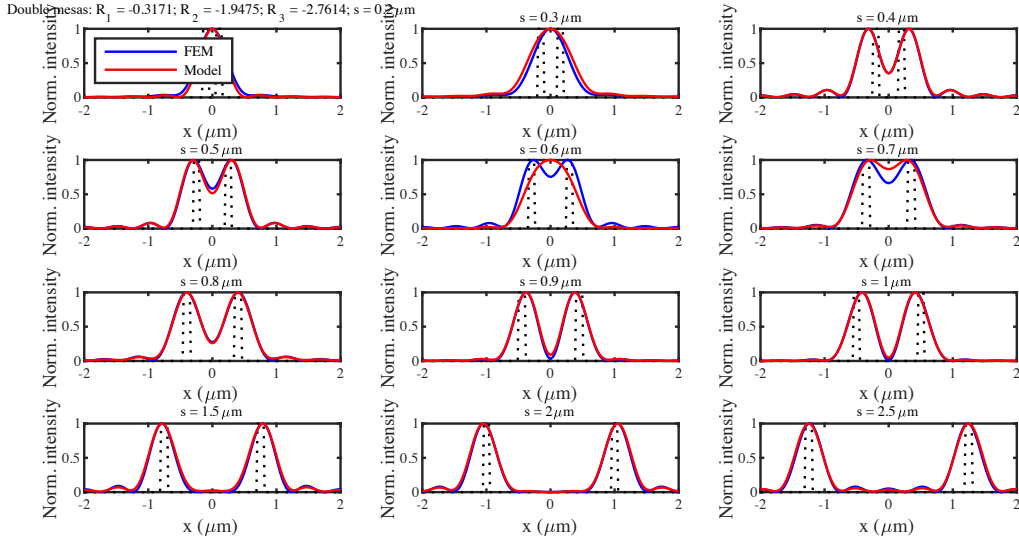


Figure 7.9: Double mesas - comparison of normalised line intensity variations predicted by the scattering model and FEM. TE polarization.

7.3 2D model – TM polarization

For TM incidence, diffraction coefficients can again be obtained from Equation 7.5. However, because secondary diffraction over mesa surfaces is significant, the corresponding diffraction coefficients (d'_3, d'_4, d'_5) and their ratios R_4 and R_5 , representing the scattering processes shown in the middle and last row of Figure 7.3a are needed. Table 7.2 shows the results for incidence at 50° to the vertical. As R_4 and R_5 are of comparable magnitude to other ratios, they must be included for TM calculations.

Numerical results obtained using these coefficients are shown in the following figures. Figure 7.10 shows the variation of the absolute amplitude along the line of imaging level of single mesas of width $w = 0.01\mu\text{m}$ to $w = 3\mu\text{m}$. For small w , the effect of surface secondary diffraction is heavily amplified due to the short distance between the two mesa corners. From the discussion in previous chapters, the magnitude of the trailing edge field

in TM incidence is stronger than the leading edge field. This is seen again in the results for mesas of width $w = 0.5\mu m$ or wider. The agreement between the simple scattering model and FEM is good, except for the narrow mesa.

d_1	d_2	d_3	d_4	d_5
+1.0344	-1.4667	+1.3067	+0.5774	+0.4143
$R_1 = d_2/d_1$		$R_1 = -1.4667/+1.0344 = -1.4179$		
$R_2 = d_3d_4/d_1$		$R_2 = +1.3067 \times +0.5774/+1.0344 = +0.7294$		
$R_3 = d_5d_4/d_1$		$R_3 = +0.4143 \times +0.5774/+1.0344 = +0.2313$		
d_3'		d_4'		d_5'
-2.6133		-1.1547		-0.8286
$R_4 = d_3'd_4'/d_1$		$R_4 = -2.6133 \times +1.1547/+1.0344 = +2.9172$		
$R_5 = d_5'd_4'/d_1$		$R_5 = -0.8286 \times -1.1547/+1.0344 = +0.9250$		

Table 7.2: TM GTD diffraction coefficients and ratios, for incidence at 50° to vertical.

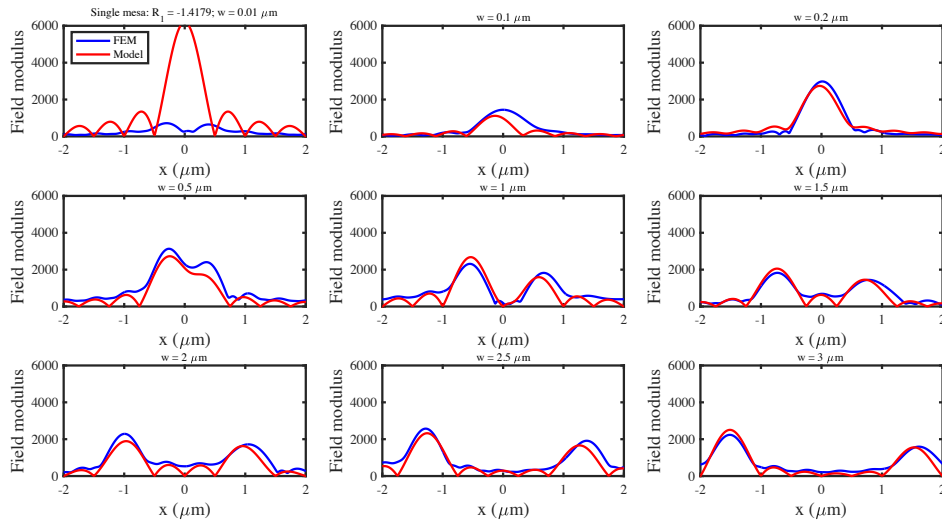


Figure 7.10: Single mesas – comparison of absolute amplitude variation along the image line as predicted by the scattering model and FEM. TM polarization.

Figure 7.11 shows a similar comparison between the line images of single trenches of different gaps $g = 0.01\mu m$ to $g = 3\mu m$ for TM incidence. For this polarization, the importance of surface secondary diffraction has been addressed many times, especially for narrow mesas. For structure like trenches, there is no surface diffraction, and secondary diffraction between two opposing corners is not as dominant as it was for single mesas. Consequently, there is better agreement between the simple scattering

model and the predictions of FEM for trenches for TM incidence than was previously demonstrated for TE incidence (Figure 7.6).

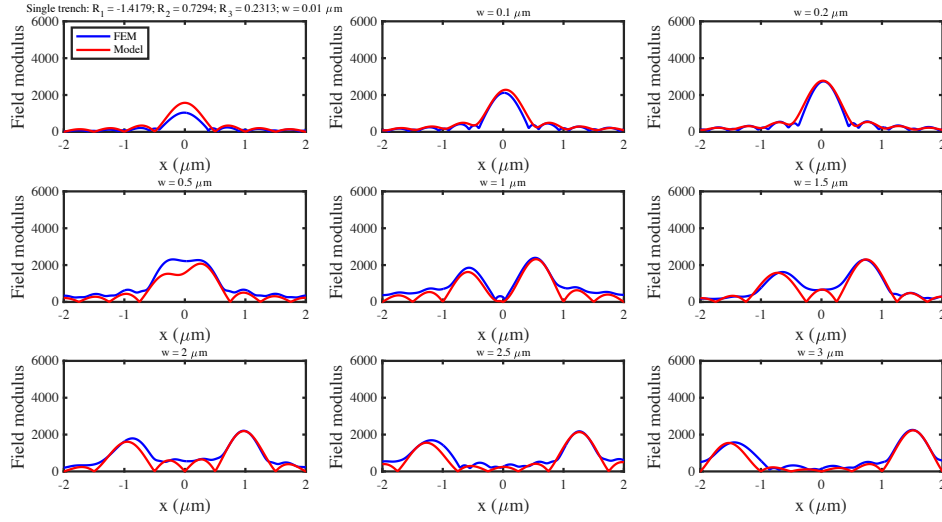


Figure 7.11: Single trenches – comparison of absolute amplitude variation along the image line as predicted by the scattering model and FEM. TM polarization.

Similarly, Figure 7.12 compares the scattering model with FEM for a double mesa configuration of width $w = 0.1\mu\text{m}$ and centre-to-centre separation $s = 0.2\mu\text{m}$ to $s = 2.5\mu\text{m}$. Since the agreement for single trenches is excellent, the double mesa model should provide similar performance. However, for $s < 1\lambda$, there are discrepancies, especially for $s = 0.6\mu\text{m}$. A possible reason is the gap between the inner sidewalls is exactly 1λ , implying some resonant effect that is not taken into consideration. However, the general matching between scattering model and FEM is still good.

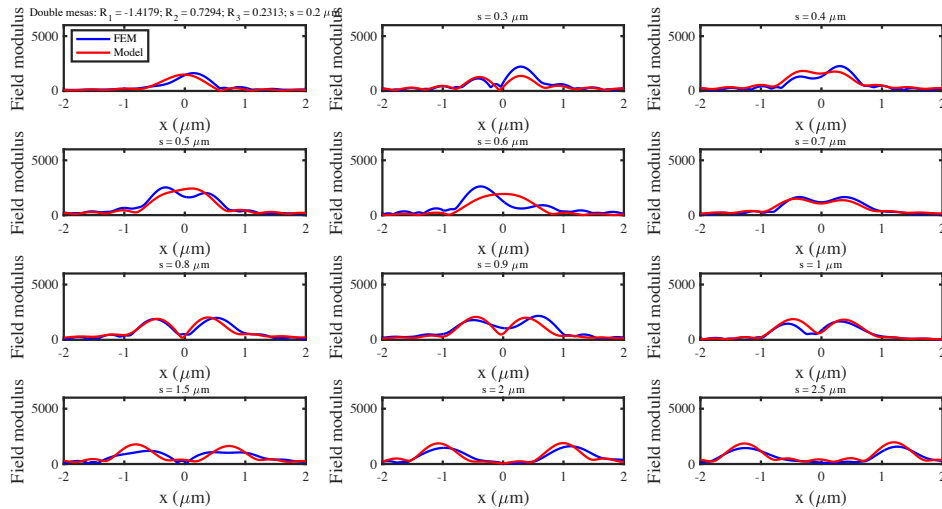


Figure 7.12: Double mesas – comparison of absolute amplitude variation along the image line as predicted by the scattering model and FEM. TM polarization.

Here the normalised intensity variations of the structures in TM incidence is omitted to avoid repetition of results, and the overall conclusion is that the simulation model can cater for TM polarization as well as TE. In the following sections, we show how it may be extended to 3D imaging.

7.4 Huygens' principle and the ITD

We first briefly review Huygens' principle, needed to convert the scattering model from 2D to 3D. Assuming that an electric field is known on a given locus or surface, this provides a simple method of estimating subsequent fields. Figure 7.13a shows a textbook example of a wave passing through an aperture. Each point in the aperture is assumed to emit a spherical wave, and the subsequent field is simply assumed as the sum of these waves. Often, only the wavefront is required, and this is constructed from tangents to arcs drawn from a set of sources. However, the field amplitude may also be found by integration. The principle was expanded by Fresnel and placed on a more rigorous footing by Kirchhoff (who included an obliquity factor that we ignore here).

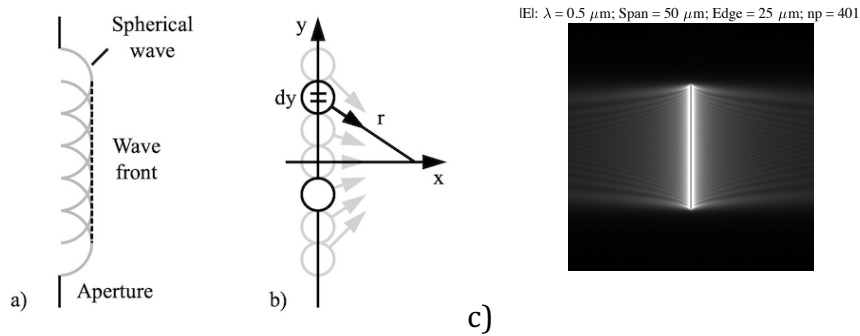


Figure 7.13: Huygens' principle: a) propagation of wavefront, b) coordinate system for field calculation; c) 2D field generated by an edge wave from a finite edge.

Huygens' principle can be applied to edge waves and can be used not only to develop an incremental theory of diffraction but also to allow the calculation of re-scattering of in-plane radiation from nearby edges. To illustrate the former, Figure 7.13b shows an infinite edge parallel to the y -axis and passing through $x = 0$, which is assumed to radiate spherical waves with amplitude A per unit length. The electric field on the x -axis can be found by summing these waves, as:

$$E(x) = \int_{-\infty}^{\infty} A \frac{\exp(-jk_0 r)}{r} dy \quad (7.15)$$

Here $k_0 = \frac{2\pi}{\lambda}$ and $r = \sqrt{x^2 + y^2}$. For large $k_0 r$, the principle of stationary phase can be invoked. In this case, only regions with $y \sim 0$ will contribute to the integral, so we may assume that y is small. We first write the radius as:

$$r = x \sqrt{1 + \frac{y^2}{x^2}} \approx x + \frac{y^2}{2x} \quad (7.16)$$

We use this approximation for the phase term but take $r \approx x$ for the amplitude term as it varies more slowly. Hence, Equation 7.15 becomes:

$$E(x) = \frac{A}{x} \exp(-jk_0 x) \int_{-\infty}^{\infty} \exp\left(-j \frac{k_0 y^2}{2x}\right) dy \quad (7.17)$$

The integral above can be evaluated analytically, giving:

$$\int_{-\infty}^{\infty} \exp\left(-j \frac{k_0 y^2}{2x}\right) dy = \sqrt{\frac{2\pi x}{k_0}} \exp\left(-j \frac{\pi}{4}\right) \quad (7.18)$$

Substituting and re-arranging, we get:

$$E(x) = A \sqrt{\frac{2\pi}{k_0}} \exp\left(-j \frac{\pi}{4}\right) \frac{\exp(-jk_0 x)}{\sqrt{x}} \quad (7.19)$$

This result is a cylindrical wave with amplitude $A' = A \sqrt{\frac{2\pi}{k_0}} \exp\left(-j \frac{\pi}{4}\right)$, so Huygens' principle implies that an infinite edge will effectively radiate a cylindrical wave.

A real edge will however be finite. In this case, the integral can be evaluated numerically. Figure 7.13c shows the variation of $|E|$ in the (x, y) plane for an edge length of $25\mu m$ and $\lambda = 0.5\mu m$, obtained by summing spherical waves at each point. Radiation is emitted laterally, and its amplitude decreases with distance on either side. Similar radiation would be expected in the (y, z) plane, implying that the scattering is at least approximately cylindrical. Some additional rapid modulation may also be seen. Figure 7.14a compares the variations of $|E(x)|$ obtained from the analytic cylindrical wave expression (blue) and numerical method (red); they are clearly similar. Figure 7.14b shows their ratio; apart from small ripples, the ratio is unity. It is simple to show that the ripples decrease as the edge lengthens, implying the radiation gradually tends to that from an infinite edge.

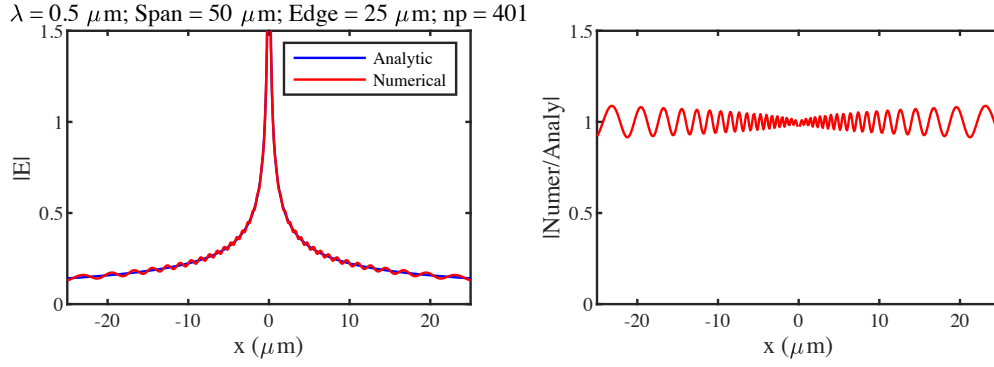


Figure 7.14: Electric field variations near an edge: a) analytic and numerical variations of $|E|$; b) ratio of numerical and analytic variations.

The generalisation offered by the incremental theory of diffraction is simply to assume that scattering from a finite or curved edge can be described in the same terms, and that each point along the edge emits a spherical wave. However, to obtain the correct correspondence, the relation between the cylindrical and spherical wave diffraction coefficients must be found. Comparison of Equation 7.19 with Equations 7.3 and 7.4 shows that equivalence is obtained if the spherical wave amplitude is given by:

$$A = E_0 \frac{d}{2\pi} \quad (7.20)$$

Thus, the spherical wave diffraction coefficient is simply $d' = d/2\pi$. There is now sufficient information to construct a 3D model of dark-field imaging.

7.5 3D model

To construct a 3D model, we combine conclusions from previous Chapters. The arguments involved can be discussed with reference to Figure 7.15, which shows dark-field illumination and imaging of simple HAR structures.

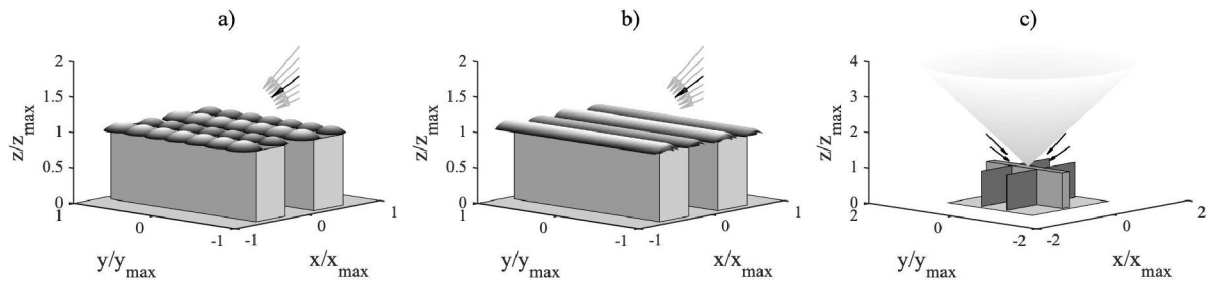


Figure 7.15: Scattering of a) spherical and b) cylindrical waves from edges; c) imaging.

We assume that:

1. Geometries are 'Manhattan-type', with mesa edges parallel to the x - and y -axes as shown in Figure 7.15a and described by 2D layouts.
2. In DF, only edge waves contribute to the image, and the locations of edges generating such waves may be found from the layout by edge detection.
3. Edge waves arise by scattering illuminating waves into spherical waves. This process yields primary out-of-plane scattering directly. However, re-scattering of in-plane radiation by nearby edges yields secondary out-of-plane scattering.
4. For a long edge, the spherical waves must sum to cylindrical waves as shown in Figure 7.15b. The spherical wave diffraction coefficient may be found from the GTD value for a cylindrical wave using Equation 7.20. However, appropriate coefficients must be assigned to leading and trailing mesa edges.
5. The number of illuminating waves can be reduced by noting that only those near normal incidence will generate an appreciable response. However, because of the slow variation with azimuthal angle, a single such wave may be sufficient. All-round illumination may then be approximated using four waves incident outside the acceptance cone of the lens, as shown in Figure 7.15c.
6. Formation of an image from the object field may be simulated by the standard route of Fourier transformation, filtering by the pupil function, and inverse transformation (again, neglecting inversion to aid in interpretation). Diffuse illumination is modelled by incoherently summing contributions from all four illuminating waves.

Implementation should make use of efficient processes such as convolution and the fast Fourier transform (FFT). The algorithms involved may conveniently be discussed with reference to Figure 7.16, which highlights the most important features.

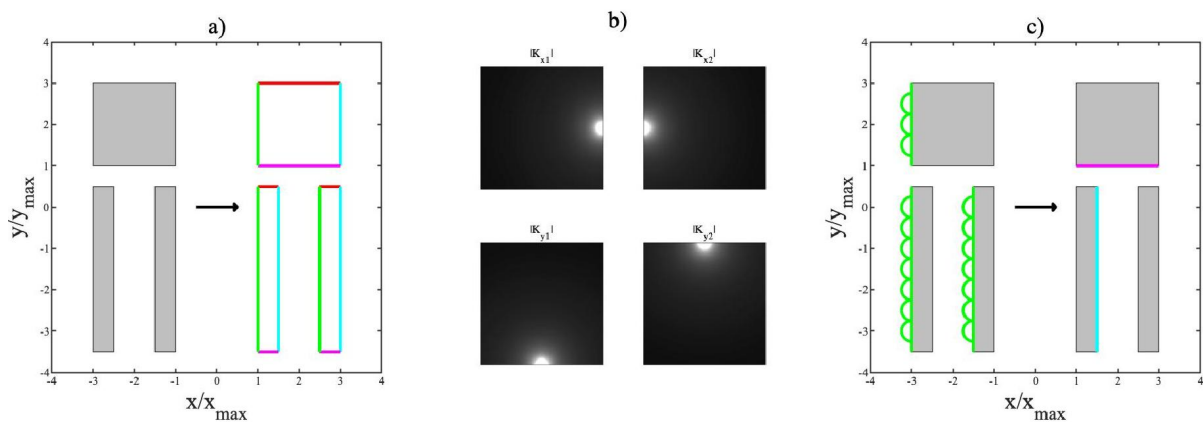


Figure 7.16: a) Edge detection and generation of primary scattering; b) kernels for in-plane radiation, and c) generation of secondary scattering from left-facing edges.

The first step is edge detection. The left-hand half of Figure 7.16a shows an example layout L containing three rectangular features marked in grey. Detection of all the edges together can be carried out by convolution, for example using a 2D Sobel filter. However, to assign the correct illumination and diffraction coefficients, the edges must be separated into four groups (left- and right-facing vertical edges, and up- and down-facing horizontal edges). Vertical and horizontal edges e_V and e_H are therefore detected separately by convolution with two different 1D Sobel filters, as:

$$e_V = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * L \quad e_H = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * L \quad (7.21)$$

Left-facing vertical edges e_{VL} are then extracted by shifting e_V to the right by one pixel, performing a pointwise AND operation with L , and shifting the result back to the left by one pixel. Right-facing vertical edges e_{VR} and up- and down-facing horizontal edges e_{HU} and e_{HD} are extracted using similar operations, to obtain the four sets of edges shown in green, cyan, red, and magenta in the right-hand half of Figure 7.16a.

Illumination patterns for each set of edges are then computed by pointwise multiplication with a complex matrix representing the appropriate wave (E_R for a right-travelling wave, and so on). For a left-facing vertical edge illuminated by a right travelling wave (for example) we then obtain $ie_{RVL} = E_R * e_{VL}$. Multiplication by the relevant diffraction coefficient (d'_1 for leading edges, and d'_2 for trailing edges) then yields the primary out-of-plane scattering (for example, as $PS_R = d'_1 ie_{RVL} + d'_3 ie_{RVR}$).

Secondary scattering is then computed as follows. Four kernels $K_{x1}, K_{x2}, K_{y1}, K_{y2}$ are constructed to describe in-plane radiation from a single point origin. Each kernel is a matrix with complex elements:

$$k_{ij} = \frac{\exp(-jk_o r_{ij})}{r_{ij}} \quad (7.22)$$

Here r_{ij} is the distance from the origin, whose location is chosen to restrict radiation to a different half-space for each kernel (to the left for K_{x1} , to the right for K_{x2} , upwards for K_{y1} and downwards for K_{y2}). The spatial variation of $|k_{ij}|$ is shown for each kernel in Figure 7.16b; the origin is located at the centre of the bright spot in each case.

Each illuminated edge pattern is then multiplied by the diffraction coefficient (d'_3 for leading edges, and d'_5 for trailing edges) and convolved with the appropriate kernel to generate an in-plane re-radiation pattern. For example, in-plane radiation from left-facing vertical edges illuminated by a wave travelling to the right is given by $\mathbf{R}_{RVL} = d'_3 \mathbf{i} \mathbf{e}_{RVL} * \mathbf{K}_{x1}$. This process effectively attaches a radiation source to every relevant edge point. The green half-circles in the left-hand half of Figure 7.16c show the result. In-plane radiation is then transferred to adjacent edges (in this case, the blue and magenta edges $\mathbf{e}_{VR}$ and $\mathbf{e}_{HD}$ in the right-hand half of Figure 7.16c) by point-wise multiplication. Multiplication by the appropriate diffraction coefficient (d'_4) then yields the secondary out-of-plane scattering, for example as $\mathbf{SS}_R = d'_4 \mathbf{R}_{RVL} * (\mathbf{e}_{VR} + \mathbf{e}_{HD})$. Primary and secondary scattering is then summed for each illuminating wave to create four different scattering patterns (for example, as $\mathbf{S}_R = \mathbf{PS}_R + \mathbf{SS}_R$).

Finally, a lens pupil function $\mathbf{P}$ is defined as a ‘top-hat’ low-pass filter. The scattering patterns from the four illuminating waves are then used to generate four images by FFT, pointwise multiplication with $\mathbf{P}$, and inverse FFT. The images are then summed incoherently to create the overall image. This algorithm was implemented in MATLAB, together with diagnostics such as extraction of line images. A variety of layouts were defined as binary matrices using a ‘case’ structure. Images of the resulting scattering were computed, and corresponding FEM images were loaded for comparison.

7.6 Numerical results

In this section, we compare simulated monochromatic images of HAR structures generated using the 3D scattering model with corresponding FEM images, mainly with TE incidence. In each case, four illuminating waves were used. As before, the optical parameters are $\lambda = 0.5\mu m$, $\theta_{max} = 30^\circ$ and $\theta_0 = 50^\circ$, the diffraction coefficient ratios are $r_1 = -0.3171$, $r_2 = -1.9475$ and $r_3 = -2.764$, and the model space was $\Lambda = 8\mu m$.

The scattering model calculations used 2001x2001 arrays, allowing 125 sample points per wavelength and suggesting that larger calculation spaces simulating more extensive devices with reduced sample density are easily possible. Unfortunately, constraints on computing resource made it difficult to provide FEM results for larger spaces, highlighting an important advantage of the scattering model. Even using this small space,

the simplest FEM simulations required 210 GB RAM and 1 hr 15 mins runtime. After exceeding the available hardware RAM (250 GB) with more complex layouts, runtimes varied from 2.5 hrs to 9 hrs. This computation time increase drastically for TM simulations as the calculation involves two components of electric field. In contrast, memory requirements and calculation times were layout-independent for the scattering model. Typical run times (including loading FEM results, and plotting images) were around 15 seconds, over $300\times$ faster.

7.6.1 Canonical objects

All the canonical objects discussed in Chapter 6 were investigated. Generally, agreement between the scattering model and FEM was excellent, but to avoid repetition, only a subset is discussed, and the most interesting features are highlighted.

Figure 7.17 shows TE results for a single vertical NEMS mesa of width $w = 0.1\mu\text{m}$. The upper panel shows the original layout (LH) and its separated edges (RH), the central panel compares simulated images obtained using the scattering model (LH) and FEM (RH), and the lower panel shows comparative line-scans across the images in the horizontal (LH) and vertical (RH) directions. The grey areas of the layout show the upper surfaces of the 3D HAR structure, and the coloured lines in the edge plot show the separate groups of extracted edges. The red and green lines in the images show the loci of line-scans. Figure 7.18 shows similar results for a single MEMS mesa of width $w = 1.0\mu\text{m}$, and Figure 7.19 shows results for a single MEMS trench of width $w = 1.0\mu\text{m}$.

In each case, the edges have been extracted correctly, and there is little to choose between the images generated using the scattering model and FEM. For single mesas, the horizontal and vertical line-scans (green) are indistinguishable. However, for the trench, the vertical line-scans (red) show some differences near the edges of the calculation window; the image intensity falls significantly here in the scattering model. This can be ascribed to the convolution method used to calculate in-plane radiation, which effectively assumes that the trench terminates at the window edge. In contrast, the periodic boundary conditions used by FEM imply that the trench continues further.

This discrepancy implies that the scattering model can simulate semi-infinite mesas

accurately, but not semi-infinite trenches. Figures 7.20 and 7.21 show further TE results of the former type, namely crossed mesas of width $w = 1.0\mu m$, and crossed double mesas of width $w = 0.1\mu m$ and separation $s = 1.0\mu m$. Here (and in subsequent results) the upper panel shows the layout (LH) and its separated edges (RH), and the lower panel compares simulated images from the scattering model (LH) and FEM (RH). Image brightness increases at the crossing points of nanoscale mesas, but not at those of microscale ones. Figures 7.22 and 7.23 show results for mesas of width $w = 0.1\mu m$ separated with a gap $g = 1.0\mu m$ (7.20) and mesa islands with a length $l = 1.0\mu m$ (7.21). In each case, there is a reduction in feature size, so that gaps appear to increase in length and islands appear to reduce. In all four cases, the scattering model accurately reproduces the FEM results.

7.6.2 NEMS devices and cellular materials

Since the original aim of this work was to provide methods of simulating optical microscope images of real NEMS, more complex layouts were investigated. Subsections of these layouts effectively contain different canonical objects.

Figures 7.24-7.27 shows TE results for an electrostatically driven NEMS resonator, whose layout is constructed in steps. Figure 7.24 shows the NEMS suspension alone, which consists of a pair of vertical nanoscale beams of width $w = 0.1\mu m$ and separation $s = 1.0\mu m$ attached to horizontal microscale anchors. As expected, the NEMS parts appear brighter than the outlines of the MEMS anchors. Figure 7.25 shows the device after the inclusion of a small, hollow suspended table formed using two additional horizontal nanoscale beams. The NEMS parts again all appear bright, and the hollow table is resolved. Figure 7.26 shows the device after the suspended table is replaced by an equivalent solid table. Removal of one nanoscale edge where the table is attached reduces the apparent brightness of the suspension beams in this region. Figure 7.27 shows the device after addition of a pair of microscale parallel-plate drive electrodes. Multiple scattering has now distorted the appearance of the nominally straight electrode edges; however, this effect appears real, since it appears in both images. There is a larger difference between the images here, arising from the small feature spacing, and closely spaced MEMS mesa edges are brighter in the FEM image.

Figures 7.28-7.31 show TE results for NEMS cellular materials based on different lattices. Figures 7.28 and 7.29 show square lattices forming a stiff mechanical metamaterial with a high equivalent Young's modulus, based on nanoscale beams of width $w = 0.1\mu\text{m}$ and separation $s = 2.0\mu\text{m}$ (7.28) and $1.0\mu\text{m}$ (7.29). In each case the four panels are as before. For the larger period, the image is as expected, namely a pattern of orthogonal lines with increased brightness at crossing points. However, for the smaller period, the image has degenerated to the crossing points, which appear as a pattern of Airy discs.

Figures 7.30 and 7.31 show TE results for weak lattices with low Young's modulus, based on nanoscale beams with width $w = 0.1\mu\text{m}$ formed into overlapping squares of side length $s = 5.0\mu\text{m}$ (7.30) and $s = 7.0\mu\text{m}$ (7.31). For the former, the proximity of the nanoscale beams in the overlap region has resulted in locally increased brightness, and there is good agreement between the scattering model and FEM. For the latter, there is a mismatch in brightness between the internal and perimeter beams obtained with the scattering model. The explanation is the lack of next-neighbour beams immediately outside the calculation window. The periodic boundary conditions assumed in the FEM model effectively provide neighbouring cells, which provide additional illumination by in-plane scattering.

7.6.3 Further limitations

We now briefly consider obvious further limitations of the scattering model. Figures 7.32 and 7.33 show TE results for solid square and cylindrical prisms of span or diameter $5\mu\text{m}$. Given the assumptions, layouts with curved edges would not be expected to be modelled correctly, and the cylinder is included for interest. In Figure 7.32, the whole of the square prism perimeter is apparently imaged correctly. However, Chapter 6 showed that limiting illumination to four waves omits scattering from external corners of square sections and from equivalent sectors of circular sections. This limitation is hidden here; not only do corners represent a detail, but both models use the same illumination. However, in Figure 7.33, it is clear: in each case, the image consists only of four bright arcs. Interestingly, the scattering model provides a reasonable approximation despite the departure from a Manhattan layout. However, the edges have clearly been extracted incorrectly, and the bright arcs are slightly asymmetric, suggesting that the Sobel filters require further investigation.

7.6.4 TM incidence

For TM incidence, the 3D images are very similar to TE images. Despite the difference in leading and trailing edges discussed in the previous sections, the predicted images from TM incidence are almost equivalent when summing up the scattered fields from different incident angles. Figure 7.34 and 7.35 show the 3D images of a suspended solid table and a resonator with TM incidence. The outlines and the features of the structure are clearly seen in the images and the image intensities are fairly constant throughout the whole structures. This is because the secondary diffraction between nearby features is negligible for TM results. It means the TM images consist mainly primary diffraction and hence the good agreement to the outlines. The TE equivalence (Figure 7.26 and 7.27) are instead a worse match due to the involvement of high degree of diffraction.

7.7 Summary

A scalar model for scattering from 2D HAR objects has been presented for DF illumination. This model avoids the complexity of full simulation by limiting the scattered field to primary edge waves and secondary edge waves generated by re-scattering of in-plane radiation. In each case, cylindrical wave diffraction coefficients are obtained from the geometric theory of diffraction. The model has been extended to 3D HAR objects by limiting features to Manhattan geometries, limiting the number of illuminating waves to four, and using Huygens' principle to calculate spherical wave diffraction coefficients for the incremental theory of diffraction. TE and TM polarizations have both been considered. The model accepts arbitrary layouts, and identification of edges and calculation of in-plane radiation are carried out by convolution with Sobel-type filters and propagation kernels. The results are in good agreement with FEM simulations, provided features are not too closely spaced, making the method suitable for simulating structures fabricated by sidewall lithography. Compared with FEM, a minimum 300-fold reduction in simulation time is achieved.

Two main additional deficiencies have been identified, associated with i) boundary effects, especially in periodic layouts, and ii) a lack of sufficient illuminating waves, especially in oblique or curved layouts. Edge effects can be mitigated by simulating a larger structure and cropping the resulting image. Different incidence angles and

wavelengths may also be included; the calculation is simply repeated for different parameters, and the results are summed incoherently. Inclusion of additional azimuthal angles is more complicated. The previous chapter showed that improved illumination of oblique edges and corners was achieved using additional illuminating waves. In principle, the only additional tools required are Sobel filters and propagation kernels that operate diagonally. Unfortunately, time constraints allowed only a preliminary investigation of these to be made, and development of an eight-wave model remains a key future goal.

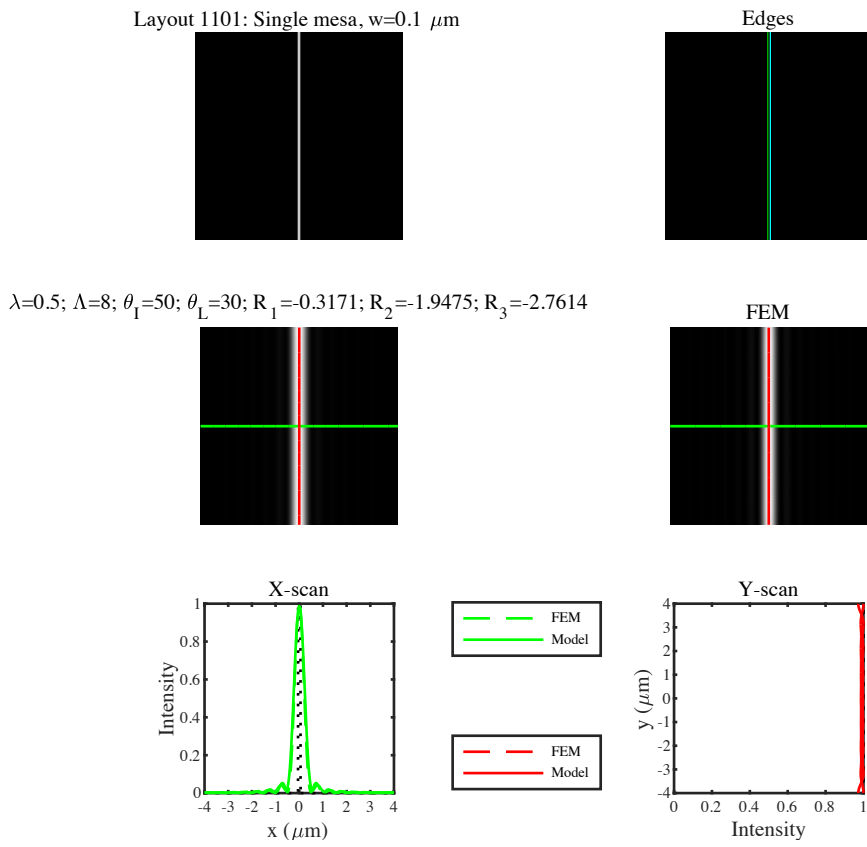


Figure 7.17: Single mesa of width $w = 0.1 \mu\text{m}$.

Upper: Layout (LH) and separated edges (RH).

Central: Simulated images obtained using the scattering model (LH) and FEM (RH).

Lower: Image line-scans obtained in the horizontal (LH) and vertical directions (RH).

TE polarization.

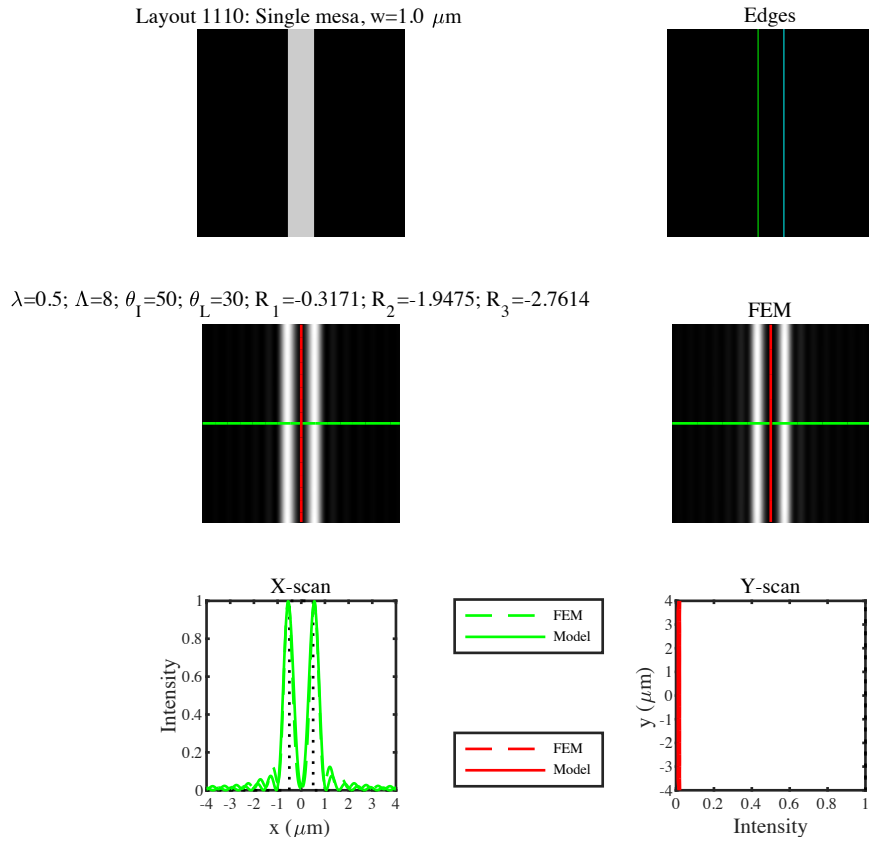


Figure 7.18: Single mesa of width $w = 1.0\mu\text{m}$.

Upper: Layout (LH) and separated edges (RH).

Central: Simulated images obtained using the scattering model (LH) and FEM (RH).

Lower: Image line-scans obtained in the horizontal (LH) and vertical directions (RH).

TE polarization.

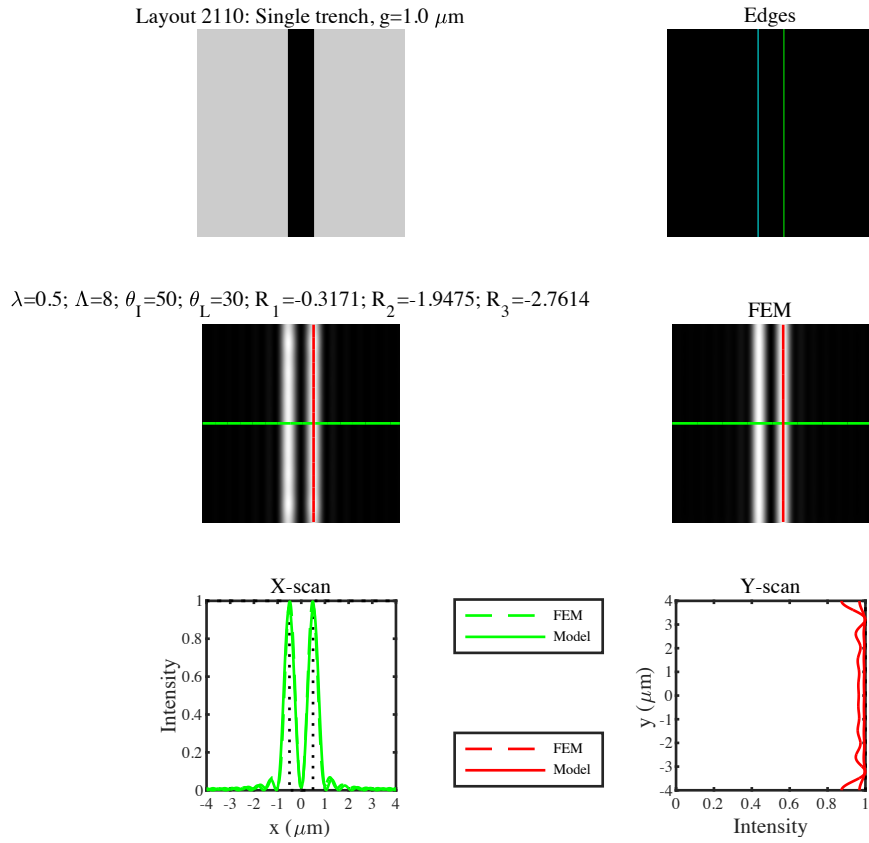


Figure 7.19: Single trench of width $w = 1.0\mu\text{m}$.

Upper: Layout (LH) and separated edges (RH).

Central: Simulated images obtained using the scattering model (LH) and FEM (RH)

Lower: Image line-scans obtained in the horizontal (LH) and vertical directions (RH).

TE polarization.

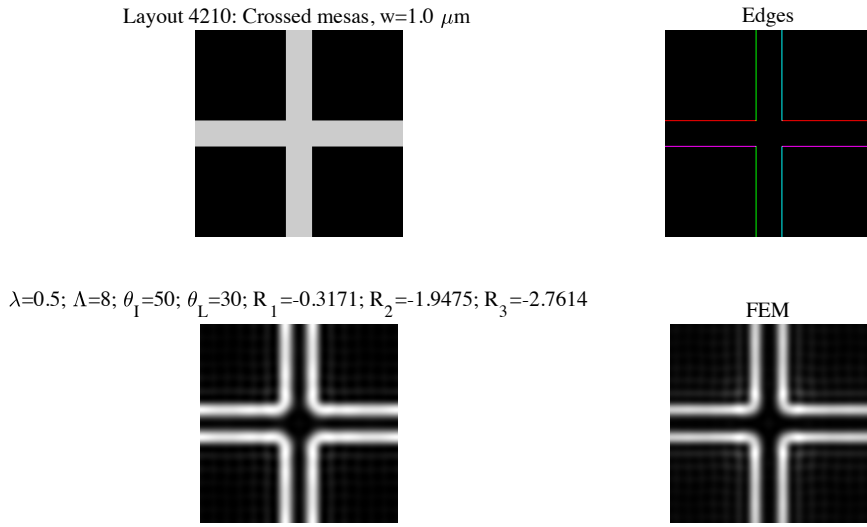


Figure 7.20: Crossed mesas of width $w = 1.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

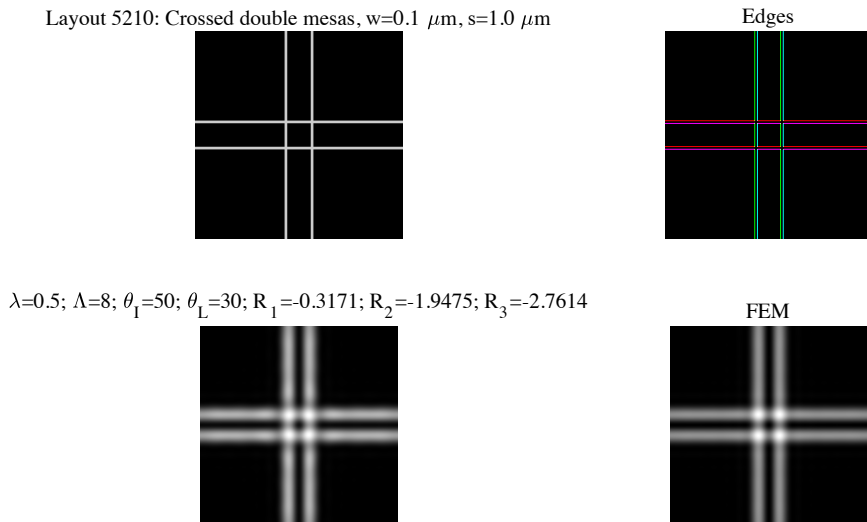


Figure 7.21: Crossed double mesas of width $w = 0.1\mu\text{m}$ and separation $s = 1.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

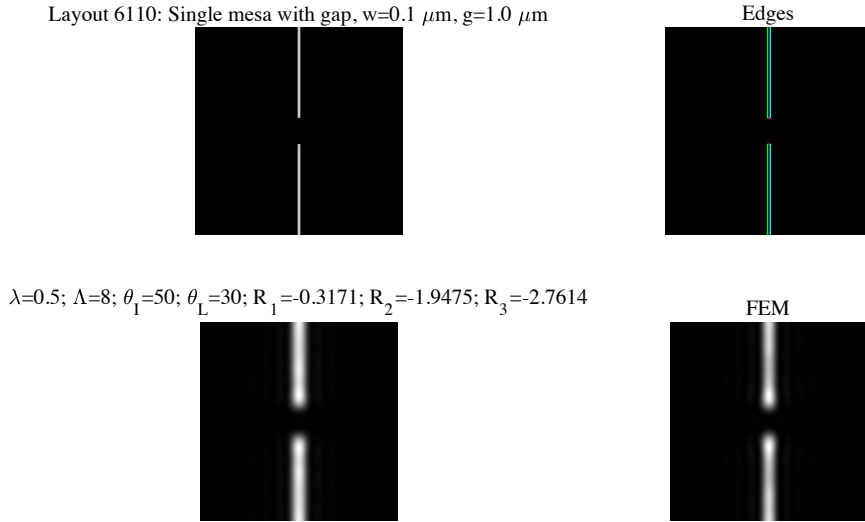


Figure 7.22: Single mesas of width $w = 0.1\mu\text{m}$ with a gap $g = 1.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

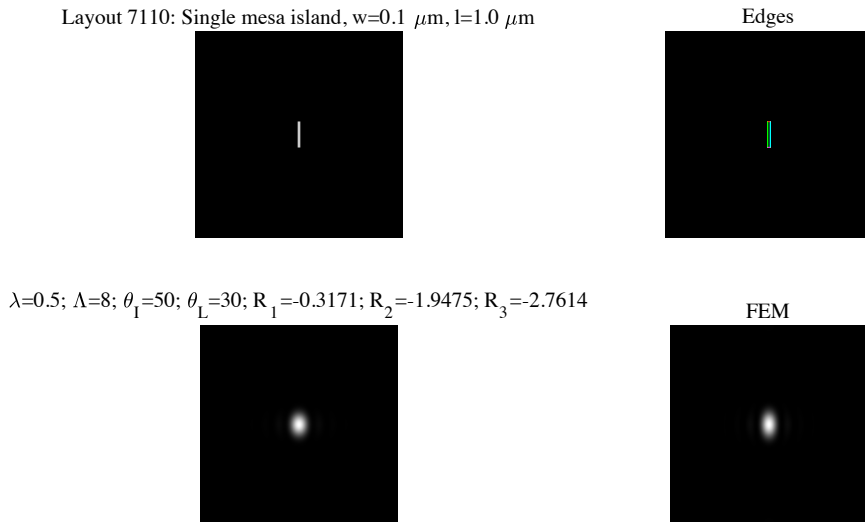


Figure 7.23: Single mesas island of width $w = 0.1\mu\text{m}$ and length $l = 1.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

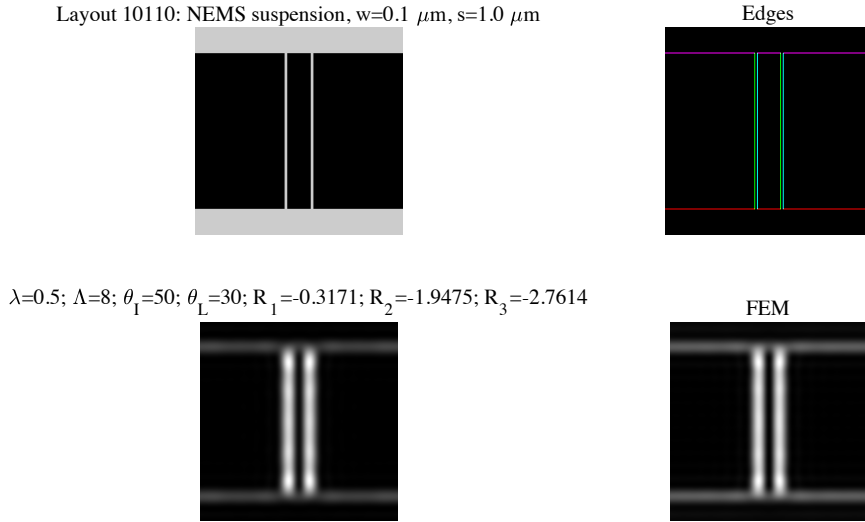


Figure 7.24: NEMS suspension of width $w = 0.1\mu\text{m}$ and separation $s = 1.0\mu\text{m}$. TE.
Upper: Layout (LH) and separated edges (RH).
Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

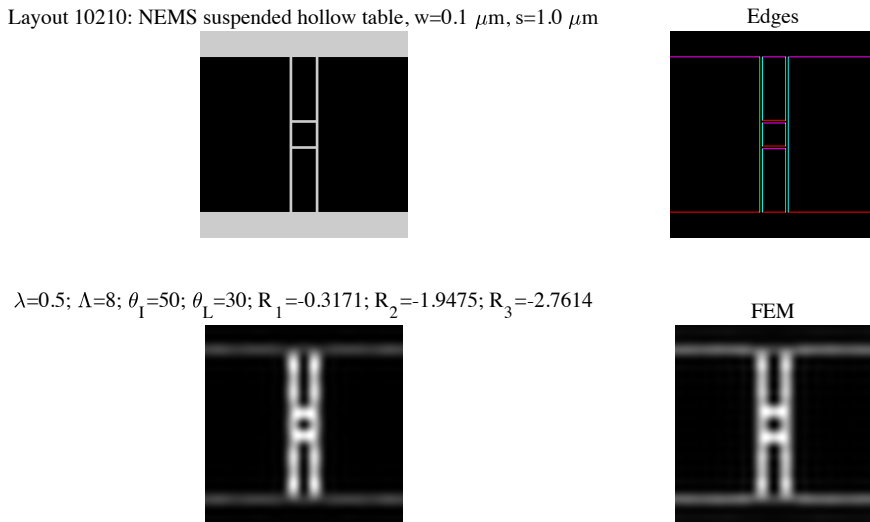
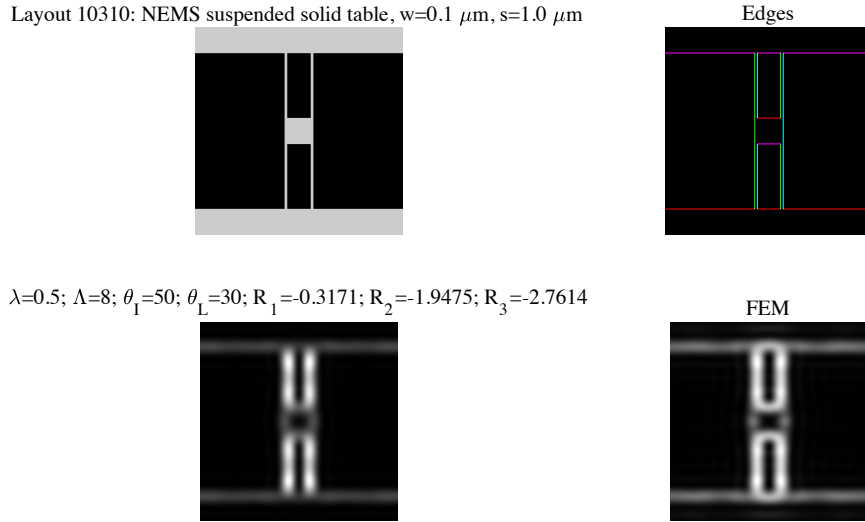
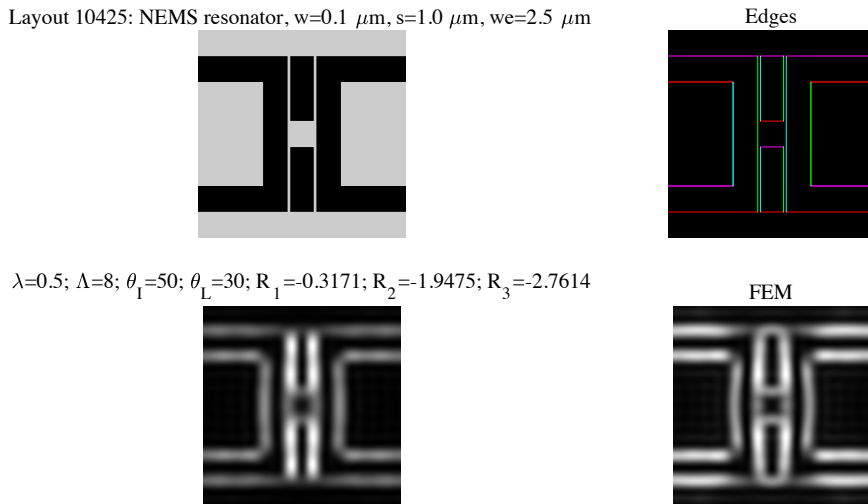


Figure 7.25: NEMS suspended hollow table of width $w = 0.1\mu\text{m}$, separation $s = 1.0\mu\text{m}$. TE.
Upper: Layout (LH) and separated edges (RH).
Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).



Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).



Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

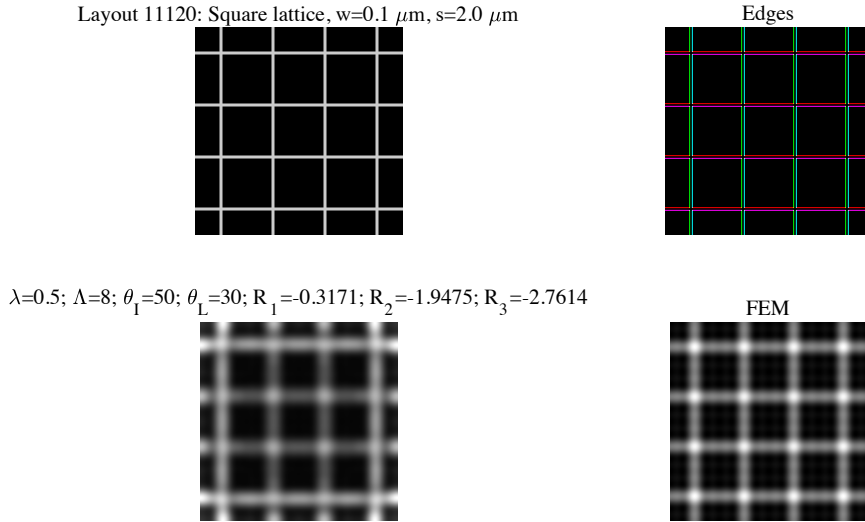


Figure 7.28: NEMS cellular lattice of width $w = 0.1\mu\text{m}$, separation $s = 2.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

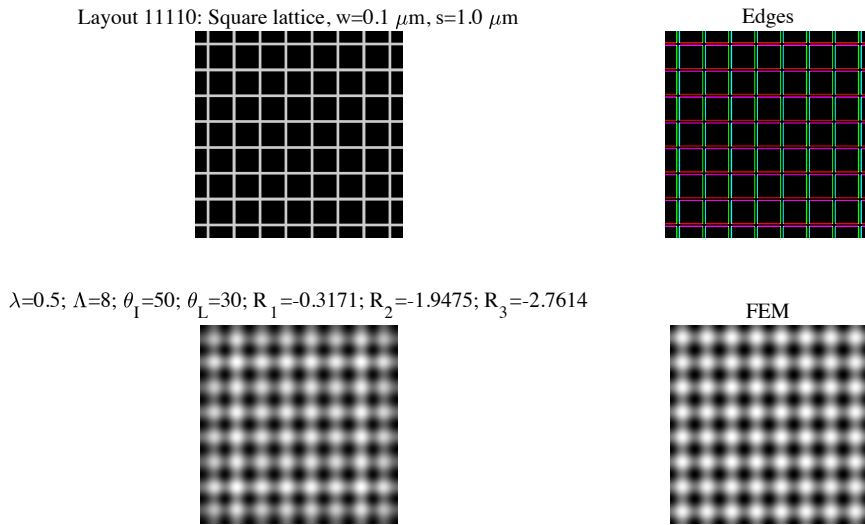


Figure 7.29: NEMS cellular lattice of width $w = 0.1\mu\text{m}$, separation $s = 1.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

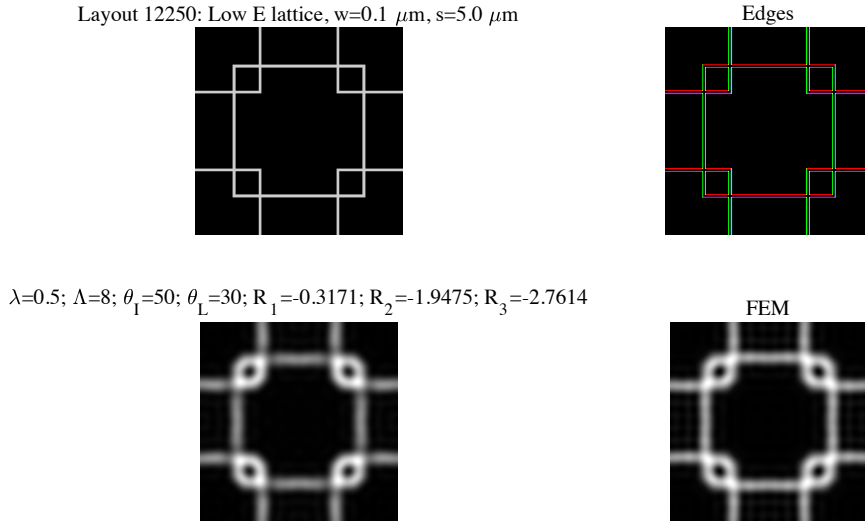


Figure 7.30: NEMS low-E lattice with width $w = 0.1\mu\text{m}$ and separation $s = 7.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

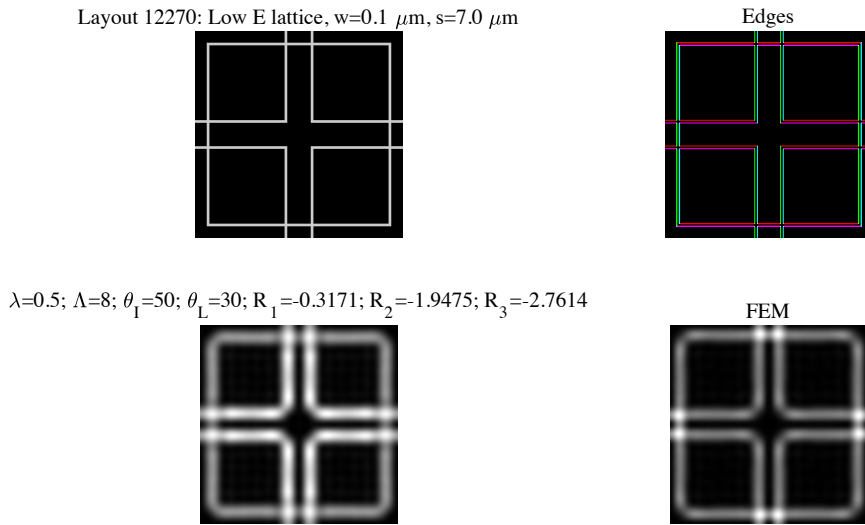


Figure 7.31: NEMS low-E lattice with width $w = 0.1\mu\text{m}$ and separation $s = 7.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

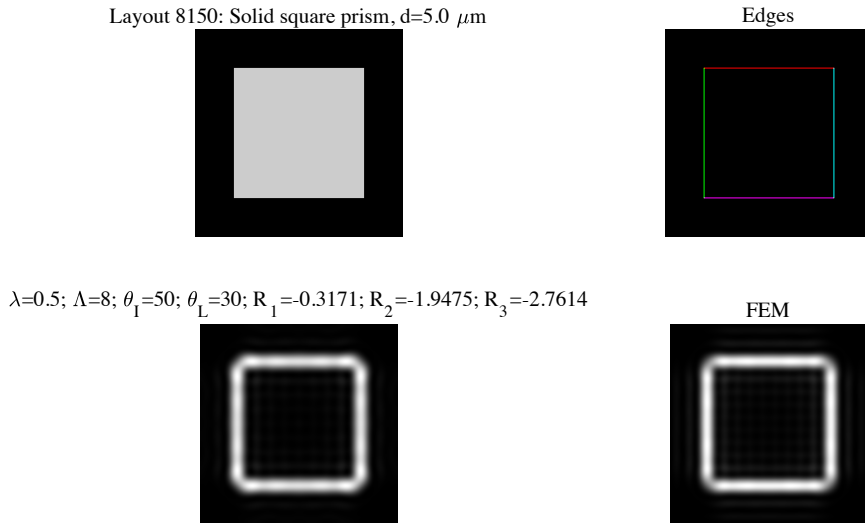


Figure 7.32: Solid square prism of span $d = 5.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

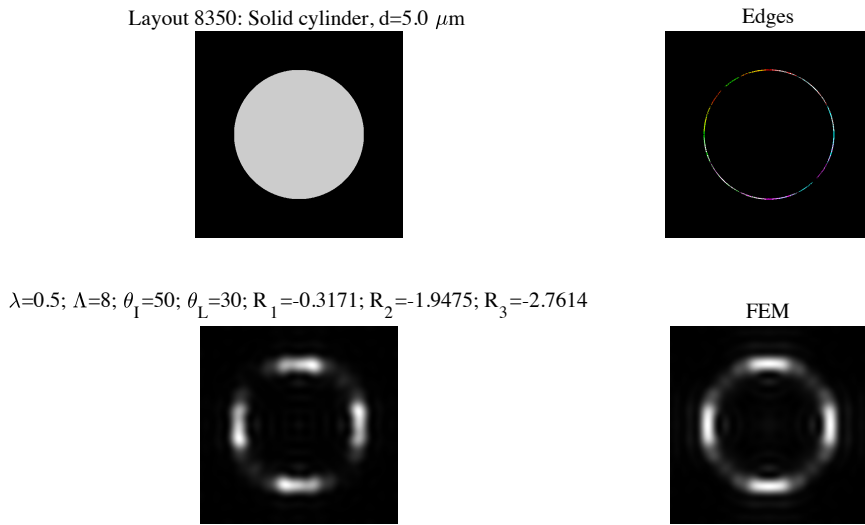
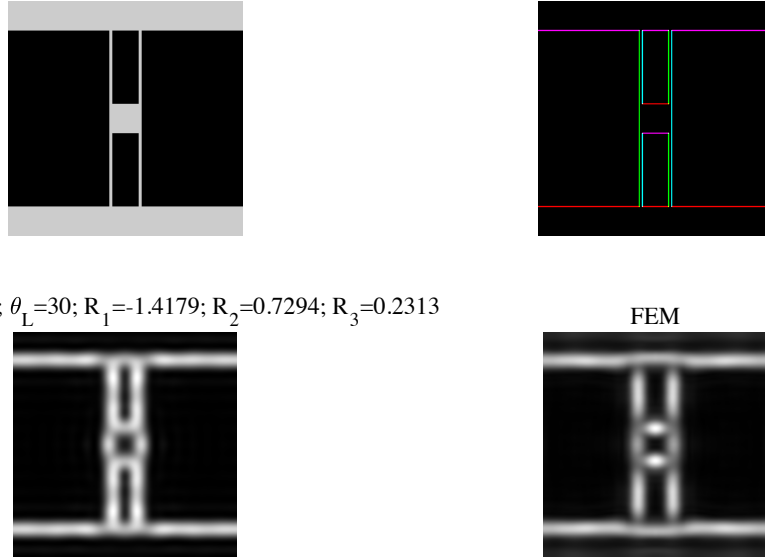


Figure 7.33: Solid cylindrical prism of diameter $d = 5.0\mu\text{m}$. TE.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

Layout 10310: NEMS suspended solid table, $w=0.1 \mu\text{m}$, $s=1.0 \mu\text{m}$



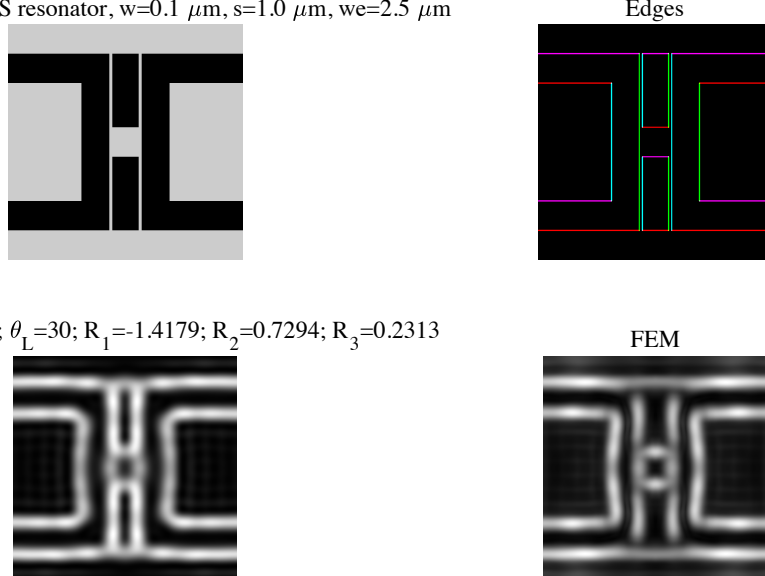
$\lambda=0.5$; $\Lambda=8$; $\theta_1=50$; $\theta_L=30$; $R_1=-1.4179$; $R_2=0.7294$; $R_3=0.2313$

Figure 7.34: NEMS suspended table of width $w = 0.1\mu\text{m}$, separation $s = 1.0\mu\text{m}$. TM.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images using the scattering model (LH) and FEM (RH).

Layout 10425: NEMS resonator, $w=0.1 \mu\text{m}$, $s=1.0 \mu\text{m}$, $w_e=2.5 \mu\text{m}$



$\lambda=0.5$; $\Lambda=8$; $\theta_1=50$; $\theta_L=30$; $R_1=-1.4179$; $R_2=0.7294$; $R_3=0.2313$

Figure 7.35: NEMS resonator of width $w = 0.1\mu\text{m}$, separation $s = 1.0\mu\text{m}$. TM.

Upper: Layout (LH) and separated edges (RH).

Lower: Simulated images obtained using the scattering model (LH) and FEM (RH).

References

Scattering from multiple knife edges

1. G. Millington, R. Hewitt, F. S. Immirzi, "Double knife-edge diffraction in field-strength predictions", *Proc. IEE – Part C: Monographs*, vol. 109, pp. 419-429, 1962.
2. J. Deygout, "Multiple knife-edge diffraction of microwaves", *IEEE Trans. Ant. Propag.*, vol. AP-14, pp. 480-489, 1966.
3. J. Deygout, "Correction factor for multiple knife-edge diffraction", *IEEE Trans. Ant. Propag.*, vol. 39, pp. 1256-1258, 1991.
4. D. S. Jones, "Double knife-edge diffraction and ray theory", *Quart. J. Mechan. Appl. Math.*, vol. 26, pp. 1-18, 1973.
5. J. Boersma, "Diffraction by two parallel half-planes", *Quart. J. Mechan. Appl. Math.*, vol. 28, pp. 405-425, 1975.
6. L. E. Vogler, "An attenuation function for multiple knife edge diffraction", *Radio Sci.*, vol. 17, pp. 1541-1546, 1982.
7. K. Furutsu, "On the theory of radio wave propagation over inhomogeneous earth", *J. Res. Nat. Bur. Stand. pt. D*, vol. 67, pp. 39-62, 1963.
8. C. L. Giovaneli, "An analysis of simplified solutions for multiple knife-edge diffraction", *IEEE Trans. Ant. Propag.*, vol. AP-32, pp. 297-301, 1984.
9. J. H. Whitteker, "Near-field ray calculation for multiple knife-edge diffraction", *Radio Sci.*, vol. 19, pp. 975-986, 1984.
10. S. Dowerah, A. Chakrabarti, "Double knife-edge diffraction with mixed boundary conditions on one of the scatterers", *J. Sound Vib.*, vol. 116, pp. 49-70, 1987.
11. P. D. Holm, "UTD diffraction coefficients for higher order wedge diffracted fields", *IEEE Trans. Ant. Propag.*, vol. 44, pp. 879-888, 1996.
12. S. A. Vavilov, M. S. Lytaev, "Modeling equation for multiple knife-edge diffraction", *IEEE Trans. Ant. Propag.*, vol. 68, pp. 3869-3877, 2020.

Scattering from double wedges

13. J. D. Hunter, R. H. T. Bates, "Secondary diffraction from close edges on perfectly conducting bodies", *Int. J. Electron.*, vol. 32, pp. 321-333, 1972.
14. R. Tiberio, R. G. Kouyoumjian, "An analysis of diffraction at edges illuminated by transition region fields", *Radio Sci.*, vol. 17, pp. 323-336, 1982.
15. R. Tiberio, R.G. Kouyoumjian, "Calculation of high-frequency diffraction by two

- nearby edges illuminated at grazing incidence" IEEE Trans. Ant. Propag., vol. AP-32, pp. 1186-1196, 1984.
16. R. Tiberio, G. Manara, G. Pelosi, R. G. Kouyoumjian, "High-frequency electromagnetic scattering of plane waves from double wedges", IEEE Trans. Ant. Propag., vol. 37, pp. 1172-1180, 1989.
 17. A. Elsherbini, M. Hamid, "Diffraction by a wide double wedge", IEEE Trans. Ant. Propag., vol. AP-32, pp. 1262-1265, 1984.
 18. A. Michaeli, "A new asymptotic high-frequency analysis of electromagnetic scattering by a pair of parallel wedges - closed form results", Radio Sci., vol. 20, pp. 1537-1548, 1985.
 19. A. Michaeli, "A hybrid asymptotic solution for the scattering by a pair of parallel perfectly conducting wedges", IEEE Trans. Ant. Propag., vol. 38, pp. 664-667, 1990.
 20. M. Schneider, R. Luebbers, "A general, uniform double wedge diffraction coefficient", IEEE Trans. Ant. Propag., vol. 39, pp. 8-14, 1991.
 21. L. P. Ivriissimtzis, R. J. Marhefka, "Double diffraction at a coplanar skewed edge configuration", Radio Sci., vol. 26, pp. 821-830, 1991.
 22. L. P. Ivriissimtzis, R. J. Marhefka, "A note on double edge diffraction for parallel wedges", IEEE Trans. Ant. Propag., vol. 39, pp. 1532-1537, 1991.
 23. F. Capolino, M. Albani, S. Maci, R. Tiberio, "Double diffraction at a pair of coplanar skew edges", IEEE Trans. Ant. Propag., vol. 45, pp. 1219-1226, 1997.
 24. M. Albani, "A uniform double diffraction coefficient for a pair of wedges in arbitrary configuration", IEEE Trans. Ant. Propag., vol. 53, pp. 702-710, 2005.
 25. G. Carluccio, F. Puggli, M. Albani, "Generalisation of UTD Double Diffraction to the case of Impenetrable Wedges with relatively general boundary conditions", IEEE Trans. Ant. Propag., vol. 62, pp. 3829-3834, 2014.
 26. G. Carluccio, M. Albani, "An efficient ray tracing algorithm for multiple straight edge diffraction", IEEE Trans. Ant. Propag., vol. 56, pp. 3534-3542, 2008.
 27. G. Carluccio, F. Pugelli, M. Albani, "A UTD triple diffraction coefficient for straight wedges in arbitrary configuration", IEEE Trans. Ant. Propag., vol. 60, pp. 5809-5817, 2012.

Incremental theory of diffraction

28. R. Tiberio, S. Maci, "An incremental theory of diffraction: scalar formulation", IEEE Trans Ant. Propag., vol. 42, pp. 600-612, 1994.

29. R. Tiberio, S. Maci, A. Toccafondi, F. Capolino, "Applications of the incremental theory of diffraction to the scattering by planar configurations", IEEE AP-S/URSI 1994, Seattle, WA, June 20-24, vol. 1, pp. 610-613, 1994.
30. R. Tiberio, S. Maci, A. Toccafondi, "An incremental theory of diffraction electromagnetic formulation", IEEE Trans. Ant. Propag., vol. 43, pp. 87-96, 1995.
31. R. Tiberio, A. Toccafondi, A. Polemi, S. Maci, "Incremental theory of diffraction: a new-improved formulation", IEEE Trans. Ant. Propag., vol. 52, pp. 2234-2243, 2004.
32. A. Toccafondi, S. Maci, A. Polemi, M. Albani, F. Capolino, "The contributions of Prof. Roberto Tiberio to the incremental theory of diffraction for electromagnetic", IEEE Ant. Propag. Soc. Int. Symp., Albuquerque, NM, July 9-14, pp. 2461-2464, 2006.

Huygens' principle

33. C. Huygens, "Traité de la lumière" van der Aa, Lieden (1690); translation by Sylvanus P. Thompson, McMillan and Co., London, 1912.
34. A. Lipson, S. Lipson, H. Lipson, "The scalar theory of diffraction", Optical Physics, Cambridge University Press, Cambridge, pp. 198-226, 2010.
35. A. J. De Witte, "Equivalence of Huygens' principle and Fermat's principle in ray geometry", Am. J. Phys., vol. 27, pp. 293-301, 1959.
36. D. Chauvat, O. Emile, M. Brunel, A. Le Floch, "Huygens' principle and Young's experiment on the propagation of light beams", Am. J. Phys., vol. 71, pp. 1196-1203, 2003.
37. W. J. R. Hoefer, "Huygens and the computer - a powerful alliance in numerical electromagnetics", Proc. IEEE, vol. 79, pp. 1459-1471, 1991.
38. P. Enders, "Huygens' principle and the modelling of propagation", Eur. J. Phys., vol. 17, pp. 226-235, 1996.

Lithography simulation using GTD

39. A. Khoh, T. Milster, B. I. Choi, G. S. Samudra, Y. H. Wu, "Influence of OPC features on the profile of 2D mask patterns", Proc. SPIE, vol. 4691, pp. 1127-1138, 2002.
40. A. Khoh, D. Flagello, T. Milster, B. I. Choi, G. S. Samudra, Y. H. Wu, "Extending a GTD-based image formation technique to EUV lithography", Proc. SPIE, vol. 5037, pp. 682-689, 2003.
41. C. B. Tan, S. H. Yeo, A. Khoh, "Modeling of wafer alignment marks using geometrical theory of diffraction", SPIE Proc., vol. 5752, pp. 977-986, 2005.
42. A. Khoh, G. S. Samudra, Y. H. Wu, T. Milster, B. I. Choi, "Image formation by use of the

- geometrical theory of diffraction", J.O.S.A. A, vol. 21, pp. 959-967, 2004.
43. G. S. Chua, C. J. Tay, C. Quan, Q. Lin, "Rigorous diffraction analysis using geometrical theory of diffraction for future mask technology", Proc. SPIE, vol. 5377, pp. 1267-1278, 2004.
 44. J. Tirapu-Azpiroz, E. Yablonivitch, "Incorporating mask topography edge diffraction in photolithography simulations", JOSA A, vol. 23, pp. 821-828, 2006.

8. Summary and Future Work

8.1 Conclusions

This work has investigated the optical imaging of high aspect ratio (HAR) nanostructures for the purpose of metrology for NEMS structures. In general, metrology is defined as the science of measurement and its application. This is particularly difficult for nano-scale structures as the current available metrology methods remain time consuming. For applications like chip inspection in manufacturing industry or device development that requires fast prototyping, imaging time and precision are crucial. Therefore, there is a need for quick inspection for nano-scale structures. This work aims to provide a possible approach for such applications. The emphasis was on understanding incoherent monochromatic imaging, and extension to polychromatic imaging is (by contrast) relatively simple. The approach taken was to simulate images using a highly simplified scattering theory adapted from rigorous diffraction theory. In contrast to most previous work, layouts may be three-dimensional and aperiodic, but are currently restricted to Manhattan geometries.

The first part of the thesis considers the scattering of electromagnetic (EM) waves from sub-wavelength structures. Previous studies have shown that NEMS devices are visible in conventional light microscopes, albeit in a modified form consisting mainly of outlines, and that the best results are obtained with dark field illumination (DF). The operation of a light microscope and scattering from 2D structures were therefore first reviewed, starting with Sommerfeld's seminal solution for diffraction from a conducting half-plane.

Different high-frequency asymptotic approaches were then considered as alternative approximations to Sommerfeld's solution. The geometrical theory of diffraction (GTD) provides approximations based on asymptotic expansions, which fail at caustic and transition regions. The uniform theory of diffraction (UTD) attempts to suppress divergent asymptotes at the shadow boundary, but near-field calculations are still erroneous. The physical theory of diffraction (PTD) approximates the caustic field correctly using surface current integration, but does not provide an explicit solution; instead, convoluted integration is required. Because optical imaging involves the far-field, GTD was therefore eventually selected as a simple basis for simulating diffraction.

The predictions of GTD for the variation of diffraction coefficient with incidence and observation angle were then compared with the results of finite element modelling (FEM) for two basic elements objects, half-planes and wedges. The results were remarkably similar, and showed that edge waves from double-sided illumination of wedges are the main contributors to DF image formation. GTD was further validated using modal diffraction theory for more general 2D periodic NEMS and a model was then developed to simulate image fields by the standard process of Fourier transformation, filtering using the pupil function, and inverse Fourier transformation. A set of canonical structures consisting of widely-spaced mesas, widely-spaced trenches and closely spaced mesas was then investigated. For 'isolated' mesas, primary scattering from edges was the important factor; re-scattering across the mesa plateau was unimportant, and the resulting image was determined mainly by the properties of the lens and the mesa width. For structures containing narrow air gaps such as trenches and closely-spaced mesas, re-scattering in gap regions is significant.

Images of more general 3D structures were then simulated, in all cases using FEM as a benchmark. The dominant effect of near perpendicular incidence was identified as a method of reducing the necessary angular spectrum of illuminating waves. DF images of canonical structures were reviewed and 3D effects (such as field superposition at feature intersections and brightness reduction at feature ends) were noted. Pairs of structures with different layouts that gave similar DF images were also identified. This effect was ascribed to the dominance of DF image formation by edge waves, so that structures with similar edges have similar images.

Because FEM simulations require extensive computer resources and run-times, a simple method of predicting dark field images was then developed. The layout is first defined in a Manhattan geometry consisting of an arbitrary set of rectangles. Convolution filtering is then used to identify edges and categorise them as leading or trailing for waves incident in just four different directions. For each incident wave, spherical edge waves are then assumed to be generated along each edge, as predicted by Huygens' principle and the incremental theory of diffraction (ITD). Diffraction coefficient values are obtained from GTD for both TE and TM polarization, allowing primary scattering to be computed directly. Secondary scattering is found by first calculating an in-plane radiation field by

convolution with a set of radiation kernels, and then re-scattering this field from nearby edges. A simulated image is generated from the complete scattered field as before. Images obtained from each illuminating wave are then summed incoherently to obtain the complete image.

Using this approach, run times were dramatically reduced by comparison with 3D FEM, and many simulated dark field images show extraordinary agreement. Viable feature widths and spacings are usefully comparable to those achievable using sidewall transfer lithography (STL) and deep reactive ion etching (DRIE), a low-cost parallel method of fabricating HAR NEMS. At this stage, the smallest size of the NEMS structure that can be measured is roughly $0.2\mu\text{m}$, that is also the smallest feature that can be fabricated with STL. This limit is improved for closely placed structure. For NEMS devices fabricated using STL, the structures can be realised with optical imaging and verified with the analytical simulation tool in this work. However, identifiable discrepancies were observed for other specific layouts, and improvements to address these are proposed in the next section. Publication of this work was published and another possible publication focusing on the indeterminate structures is currently being discussed with my supervisors.

8.2 Future work

The scattering model currently assumes illumination of symmetric Manhattan geometries with large spacing at just four azimuthal angles, a single out-of-plane angle (50°), and a single wavelength ($0.5\mu\text{m}$). Despite the excellent results achieved using this drastic simplification, the model could easily be extended and validated as follows:

8.2.1 Model improvements

Extension to asymmetric structures is trivial; symmetry was simply assumed to halve FEM run-times. Inclusion of additional out-of-plane angles and wavelengths is also simple, and merely requires incoherent summation of results obtained by repetitive use of existing algorithms. Inaccurate modelling of re-scattered fields near the edges of calculation windows could also easily be managed by simulating larger structures and cropping the resulting simulated images. However, accurate modelling of small feature spacing requires development of fast methods for computing multiple scattering of in-

plane radiation. This may be a tedious task, due to the likely slow convergence of the overall sum and the difficulty of extending a suitable 2D algorithm to 3D.

Extension to curved layouts and additional azimuthal angles is also difficult, because edges must first be identified at arbitrary angles and then illuminated by near-normal waves. However, as noted earlier, extension to eight-sided polygons illuminated at eight azimuthal angles should be straightforward, and can be implemented by developing convolution operators for edge detection and re-radiation that operate diagonally. The effect of defocus and/or variable feature height could be included using appropriate phase shifts in the Fourier or spatial domains, and the effect of a substrate could also be included using image sources. Finally, the effect of imperfect conductors such as silicon could be described by including material factors in the diffraction coefficients.

8.2.2 Comparison with experiment

At the start of this project, a key aim was to compare simulated images with experimental results from NEMS devices. Some experimental images were previously shown in Chapter 1. An obvious challenge is to verify these against exactly corresponding structures. Unfortunately, the coronavirus pandemic resulted in a 2-year closure of the Departmental clean rooms and the departure of some associated fabrication expertise. Available samples are only approximately corresponding, and available images are also polychromatic.

However, Figure 8.1 shows a comparison between simulated and experimental DF images of a cellular NEMS metamaterial with low Young's modulus. Figure 8.1a shows the predictions of the scattering model (LH) and FEM (RH) in TE-polarization, and Figure 8.1b is the TM-polarization. Figure 8.1c shows the predicted unpolarized image of the structure as the sum of the TE and TM images predicted by the imaging tool. Figure 8.1d shows an image of a structure formed in bonded silicon-on-insulator material by STL and DRIE. NEMS part widths are $0.1\ \mu\text{m}$, and the etch depth is $4\ \mu\text{m}$. The main difference is the cell size ($8\ \mu\text{m}$ for the simulated images, $16\ \mu\text{m}$ for the real structure) due to the limitation of computing resources. Also, the fabrication of the cellular material using multiple STL leaves an uneven feature depth throughout the whole structure. Despite such structural differences, the prediction and the experimental images agree

qualitatively, particularly for the extra brightness for the interconnection between each cells. This validates many of the assumptions of this project. Work of this type could easily be continued.

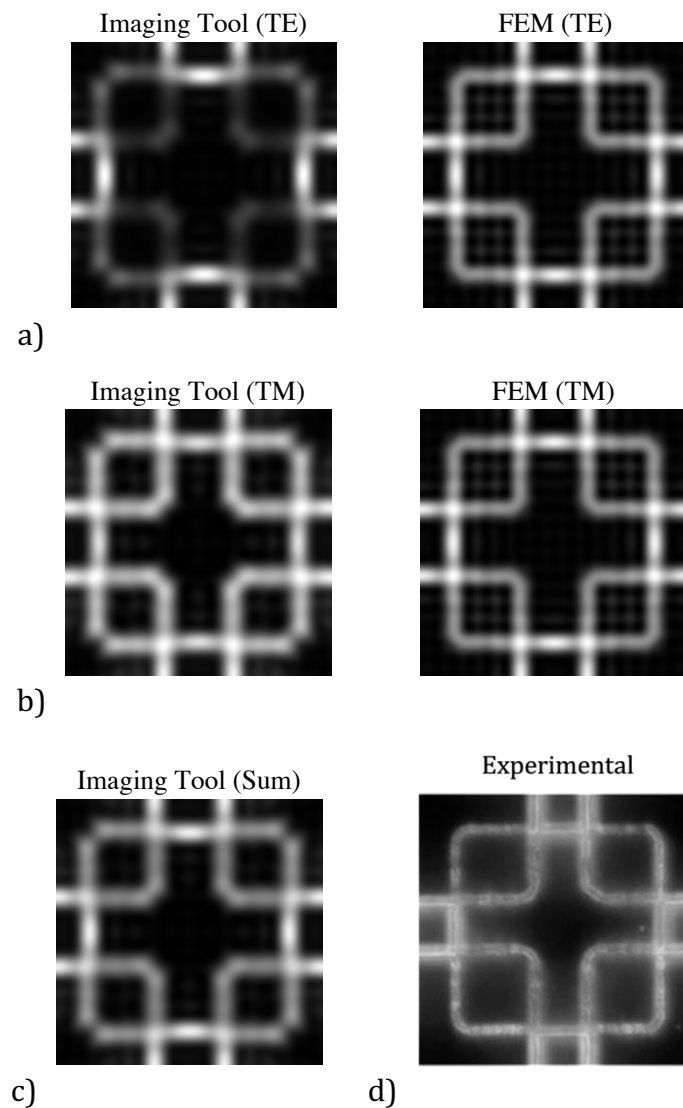


Figure 8.1 DF images of a cellular NEMS material predicted by the imaging tool and simulated from numerical calculation in a) TE polarization; and b) TM polarization. c) B/W equivalent polychromatic dark-field images of a cellular NEMS material; d) approximately corresponding experimental image.